

ISSUE TWO

NewScientist THE COLLECTION

THE UNKNOWN UNIVERSE

WHERE DID IT ALL COME FROM?

WHY IS REALITY LIKE IT IS?

DO DARK MATTER AND ENERGY EXIST?

WHAT ARE BLACK HOLES?

IS TIME AN ILLUSION?

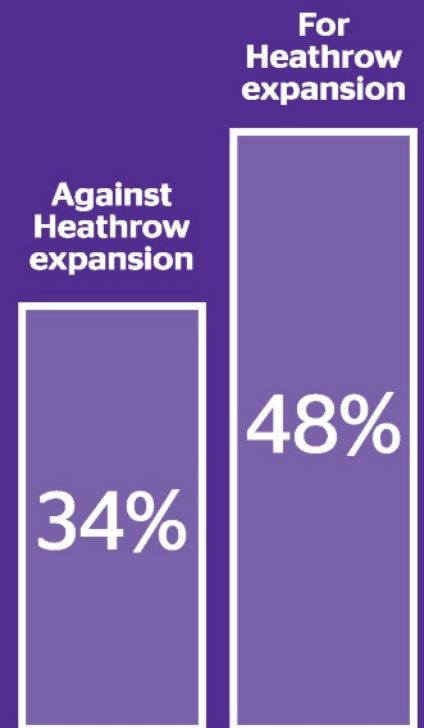
WILL WE EVER HAVE A THEORY OF EVERYTHING?

HOW WILL IT END?

£9.99



Contrary to
popular belief,
backing Heathrow
won't
make you
unpopular.



In a recent Populus poll*, nearly half of west London residents said they supported expansion at Heathrow, with just over a third opposed. In addition, 57% of voters said they felt positive towards Heathrow (just 6% were negative). The truth is, most local people appreciate the benefits that living near the UK's only hub airport brings to their communities, their families and the nation as a whole.

Heathrow

*Populus interviewed 1,000 adult residents in seven Boroughs local to Heathrow Airport between February and March 2014. In total, 7,000 residents were interviewed. Results were weighted to be demographically representative of all adults in each borough.

ISSUE TWO
**THE
UNKNOWN
UNIVERSE**

**NEW SCIENTIST
THE COLLECTION**

Lacon House, 84 Theobald's
Road, London WC1X 8NS
+44 (0)20 7611 1202
enquiries@newscientist.com

Editor-in-chief Graham Lawton
Editor Michael Brooks
Art editor Craig Mackie
Picture editor Adam Goff
Subeditor Richard Lim
Graphics Nigel Hawtin
Production editor Mick O'Hare
Project manager Henry Gomm
Publisher John MacFarlane

© 2014 Reed Business
Information Ltd, England
New Scientist The Collection is
published four times per year by
Reed Business Information Ltd
ISSN 2054-6386

Printed in England by Polestar
(Bicester) and distributed by
Marketforce UK Ltd
+44(0)20 3148 3333
Display advertising
+44(0)20 7611 1291
displayads@newscientist.com

The sky is no longer the limit

Of the 100 billion humans that have ever walked the Earth, very few will have lived their lives without wondering what is up there beyond the clouds. Every culture has constructed a cosmology – an explanation for what the universe contains and how it came to be.

For us, though, cosmology is different. We are lucky enough to be alive at a time when such questions are being answered not through storytelling, conjecture or religious tradition, but through gathering of evidence that can be subjected to rigorous mathematical analysis and experimental testing. In other words, science.

As a result, cosmology is finally coming of age. There is still a lot that we don't know, but compared with our ancestors we are incredibly knowledgeable. We are pretty sure, for instance, that the universe is just under 14 billion years old. Considering that all we have to go on is what can be seen from Earth through telescopes and the power of the human brain, that is an astonishing thing for us to have discovered.

This second issue of *New Scientist: The Collection* is dedicated to the amazing science of cosmology. A compilation of classic articles from *New Scientist*, it explores what we know – and still don't know – about the birth, life and death of the universe.

We open, appropriately enough, with our understanding of the early universe. Looking back 14 billion years is no mean feat, but cosmologists are slowly building a detailed picture of the turbulent dawn of the cosmos.

In Chapter 2 we go even deeper, exploring what lies beneath our everyday conceptions of space and time, and where those four dimensions of reality came from. We ask whether there is more than one universe, and how the bizarre world of quantum theory fits into our conceptions of reality.

Chapter 3 gets down to the nuts and bolts: what space is, and how it has been shaped throughout its existence by gravitational waves, magnetic fields and maybe some other forces that we aren't yet aware of.

Chapter 4 tackles perhaps the biggest puzzle of them all: why we have a grasp on only 4 per cent of the cosmos. The largest component of the universe is dark energy, but we know next to nothing about it. The next largest, dark matter, has been in our sights for 80 years now – are we any closer to an explanation of its composition?

Chapter 5 gets even darker. Not only are black holes a possible explanation for dark matter, but they may also hold the key to understanding how everything in the universe holds together.

In Chapter 6, we turn to one of the deepest mysteries of all: the nature of time. Einstein threw our common sense ideas of it away when he invented relativity. Can we rescue time, or must we abandon it as an illusion created by our brains?

Finally, we look to the future. Chapter 7 explores the new paths cosmology is following in pursuit of a theory of everything – and our best guesses as to how the universe will end. ■

Michael Brooks, Editor

CONTENTS

ISSUE TWO THE UNKNOWN UNIVERSE

NewScientist
THE COLLECTION

CONTRIBUTORS

Robert Adler

is a writer based in California and Mexico

Anil Ananthaswamy

is a consultant for *New Scientist*

Stephen Battersby

is a consultant for *New Scientist*

Michael Brooks

is a consultant for *New Scientist*

Matthew Chalmers

is a freelance writer based in Bristol, UK

Marcus Chown

is a consultant for *New Scientist*

Stuart Clark

is a consultant for *New Scientist*

Amanda Geffer

is a consultant for *New Scientist*

Lisa Grossman

is a reporter for *New Scientist*

Abraham Loeb

is director of the Institute for Theory and Computation and chair of the astronomy department at Harvard University

Maggie McKee

is a science writer based in Boston, Massachusetts

Katia Moskvitch

is a science writer based in London

Andrew Pontzen

is a cosmologist at University College London

Jonathan Pritchard

is a lecturer in astrostatistics at Imperial College London

Michael Slezak

is *New Scientist's* Australasian reporter

Vlatko Vedral

is professor of quantum information theory at the University of Oxford and the Centre for Quantum Technologies in Singapore

Cover image

Simon Danaher

The articles in this collection were first published in *New Scientist* between 2005 and 2013. They have been fully updated and revised.



1

The early universe

- 6 In the beginning
- 10 Bang goes the theory
- 16 Pop-up universe
- 20 The missing reel

2

The nature of reality

- 24 Quantum shadows
- 28 Seeing triple
- 32 The many faces of the multiverse

3

The fabric of the cosmos

- 38 Ear on the universe
- 42 Space against time
- 46 Beyond space-time
- 50 Warning light
- 54 The forgotten force



4

Dark stuff

- 58 Shaken and stirred
- 62 Black hole universe
- 66 Portraits of darkness
- 70 Chasing shadows
- 74 Pulling power



5

Black holes

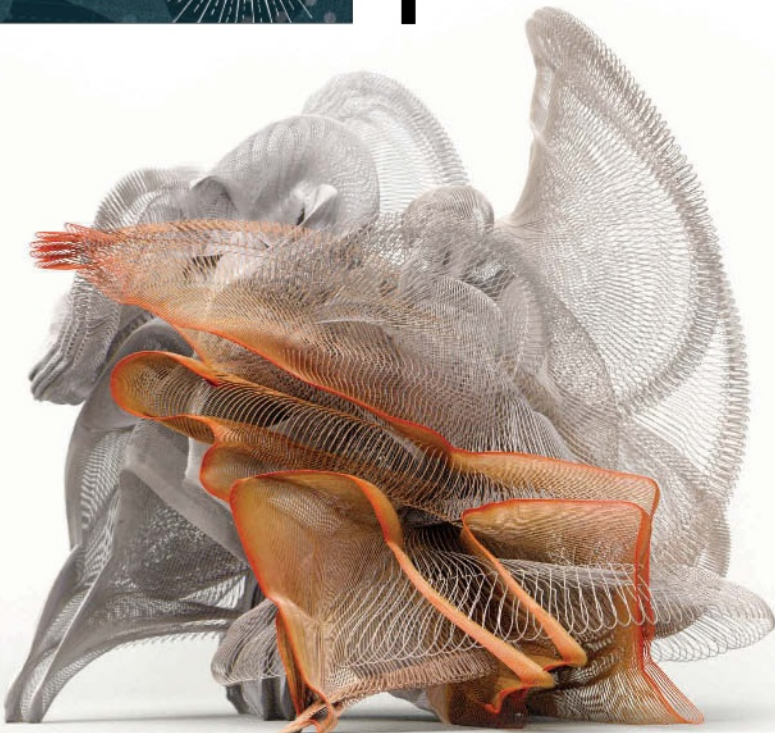
- 78 Monster mix
- 83 Black bombshells
- 86 Ring of fire
- 90 Firefight at the black hole corral



6

Time

- 92 The now delusion



7

New directions

- 97 Are we nearly there yet?
- 100 In from the cold
- 106 The infinity illusion
- 110 Ghost matter
- 116 Differently equal
- 122 Reality bits
- 126 The grand finale





Fuel consumption for the 14MY XF range in mpg (l/100km): Urban 16.7-48.7 (16.9-5.8); Extra Urban 32.8-64.2 (8.6-4.4); Combined 24.4-57.7 (11.6-4.9); CO₂ Emissions 270-129 g/km. Fuel consumption for the Jaguar XJ range in mpg (l/100km): Urban 16.7-39.8 (16.9-7.1); Extra Urban 32.8-51.4 (8.6-5.5); Combined 24.4-46.3 (11.6-6.1). CO₂ Emissions 270-159 g/km.

The figures provided are as a result of official manufacturer's tests in accordance with EU legislation. A vehicle's actual fuel consumption may differ from that achieved in such tests and these figures are for comparative purposes only.

**THE BEST THING
ABOUT WORKING LATE?
THE DRIVE HOME.**



**THE JAGUAR BUSINESS RANGE.
FROM 129 G/KM CO₂.**

Your teams aren't like the rest. And neither is Jaguar. The Jaguar business range offers something unique. An exceptional drive and high standard specification, yet excellent fuel economy and low monthly costs. A company car like no other, with the support of an award-winning network of specialist fleet & business dealers. Jaguar mean business. Do you?

XF SPORTBRAKE • XF SALOON • XJ

**WWW.JAGUAR.CO.UK/BUSINESS
0845 600 2214**

HOW ALIVE ARE YOU?





IN THE BEGINNING

Has the cosmos existed forever, or did something bring it into existence? **Marcus Chown** grapples with the universe's original mystery

AS BIG questions go, it's hard to beat. Has the universe existed forever? Over the years, some of the greatest minds in physics have argued that no matter how far back in time you go, the universe has always been here. Others have argued that the opposite must be true – something must have happened to bring the cosmos into existence. With both sides claiming that observations support their view, until recently an answer seemed as distant as ever.

However, cosmologists Alex Vilenkin and Audrey Mithani claimed to have settled the debate. They have uncovered reasons why the universe cannot have existed forever. Yet what nature grudgingly gives with one hand, it takes back with the other – even though the universe has a beginning, its origins may be lost in the mists of time.

Modern cosmology began in 1916 when Einstein applied his newly formulated theory of gravity, general relativity, to the biggest gravitating mass he could think of: the entire universe. Like Newton, Einstein favoured an unchanging universe – one that had existed forever and therefore had no beginning. To achieve this, Einstein realised, the gravity pulling together all the matter in the universe had to be countered by a weird cosmic repulsion of empty space.

Einstein's universe was, unfortunately, unstable. As the physicist Arthur Eddington

pointed out, it was balanced on a knife-edge between runaway expansion and runaway contraction. A further blow came in 1929 when astronomer Edwin Hubble observed that galaxies were flying apart from each other like pieces of cosmic shrapnel. The conclusion was that the universe was expanding.

Yet if the universe was expanding, an unavoidable consequence must be that it had been smaller in the past. Imagine rewinding that expansion back to a time when everything was compressed into the tiniest of volumes. This was the big bang.

The big bang theory that has subsequently developed describes the evolution of the universe from a hot, dense state, but it does not say anything about what brought the universe into existence. That still leaves crucial questions unanswered – what happened before the big bang, and was there really a beginning?

Little wonder, then, that the appeal of the eternal universe became popular, not least because such awkward questions need never be asked. In 1948, Fred Hoyle, Hermann Bondi and Tommy Gold proposed that, as the universe expanded, new matter fountained into existence in the gaps between galaxies and coalesced into new galaxies.

According to this steady-state picture, the universe looks the same today as it has always done and always will. It has no beginning; it has simply existed forever. ➤

However the steady-state theory was scuppered by two observations. The first was the discovery in the 1960s that the distant, and therefore early, universe does not look the same as today's universe. The second was the discovery in 1964 of the cosmic microwave background, the hot afterglow of the big bang fireball. More recently, the European Space Agency's Planck telescope has made detailed measurements of this cosmic microwave background and shown that the big bang happened 13.8 billion years ago.

Flaws in forever

A further blow to the eternal universe came from theory. In the 1960s, Roger Penrose and Stephen Hawking were two young theorists at the University of Cambridge. Their work showed that if you reversed the expansion of the universe, it is impossible to avoid reaching a point known as a singularity, where physical parameters such as density and temperature skyrocket to infinity. Crucially, physics breaks down at a singularity, making it impossible to predict what lies on the other side. According to Penrose and Hawking, the big bang must truly be the beginning.

So, story over? Well, no. It turns out that there is a loophole in the singularity theorems of Penrose and Hawking. According to Newton's laws, the gravitational pull of an object depends only on its mass. Einstein's insight showed that the strength of gravity also depends on an object's energy density and, crucially, its pressure. In deriving their powerful theorems, Penrose and Hawking had assumed that the pressure of space is always small and positive. But what if they were wrong? "It is just this possibility that has opened the way to modern cosmological theories in which the big bang is not a beginning at all," says Vilenkin. "Chief among them is inflation."

Inflation, a theory that Vilenkin helped to create, starts with a vacuum in an unusually high energy state and with a negative pressure. Together these give the vacuum repulsive gravity that pushes things apart rather than drawing them together. This inflates the vacuum, making it more repulsive, which causes it to inflate even faster.

But the inflationary vacuum is quantum

in nature, which makes it unstable. All over it, and at random, bits decay into a normal, everyday vacuum. Imagine the vacuum as a vast ocean of boiling water, with bubbles forming and expanding across its length and breadth. The energy of the inflationary vacuum has to go somewhere, and it goes into creating matter and heating it to a ferocious temperature inside each bubble. It goes into creating big bangs. Our universe is inside one such bubble that appeared in a big bang 13.8 billion years ago.

One of the striking features of inflation is that it is eternal. New high-energy vacuum is created far faster than it is eaten away by its decay into ordinary vacuum, which means that once inflation starts, it never stops and universes bubble off forever in the future. But because eternal inflation avoids the dreaded singularity, it opens up the possibility that this has always been the case with universes bubbling off forever in the past too.

Inflation is compatible with all our observations and Vilenkin is fairly certain it is fundamentally correct. Yet there is a problem with eternal inflation, which Vilenkin discovered in 2003 when he teamed up with Arvind Borde of Long Island University in New York and inflation pioneer Alan Guth of the Massachusetts Institute of Technology.

They calculated what would happen in a growing universe and made no assumptions about energy or gravity. Their theorem simply assumed that on average the universe expands. "To our amazement, it showed that space-time does not continue forever in most past directions," says Vilenkin. "Inflation must have a beginning."

However inflation is not the only game in town. So could the alternative scenarios have a beginning? In 2012, Vilenkin teamed up with Audrey Mithani, his colleague at Tufts University in Medford, Massachusetts, to examine two of the leading alternative cosmological scenarios.

The first is the "cyclic universe" developed within string theory by Neil Turok of Canada's Perimeter Institute for Theoretical Physics in Waterloo, Ontario, Canada, and Paul Steinhardt of Princeton University. In this scenario, our universe is a four-dimensional island, or "brane", in a higher dimensional space. It collides repeatedly with a second



brane (see diagram, right). Think of the two branes as two parallel slices of bread, coming together along a fifth dimension, passing through each other, pulling apart again, then coming together again. Each time the branes touch, their tremendous energy of motion along the fifth dimension creates matter on each brane and heats it to tremendous temperature. To observers on the brane, it looks exactly like a big bang and would lead to the same patterns in the cosmic microwave background and distributions of galaxies. Yet it is a big bang without a beginning, say Turok and Steinhardt, because the cycles have been repeating for eternity.

However, Vilenkin and Mithani have shown that the cyclic universe cannot continue indefinitely towards the future and the past. According to the theory, matter on the branes expands more with each cycle and this means that the Borde-Guth-Vilenkin theorem of there being a beginning to the universe still applies. "If you run it backwards like a movie in reverse, the cyclic universe encounters either a singularity or some kind of beginning like inflation," he says.

Another cosmological scenario considered by Vilenkin and Mithani is even weirder than the cyclic universe and inflation. This is the "emergent universe" imagined by George Ellis of the University of Cape Town in South Africa and Roy Maartens of the University of Portsmouth, UK. It begins as a small static universe, which exists in this state for an infinite amount of time before suddenly being

"For the first time in history we have the tools to address the origin question scientifically. I have a feeling we are getting near the truth"

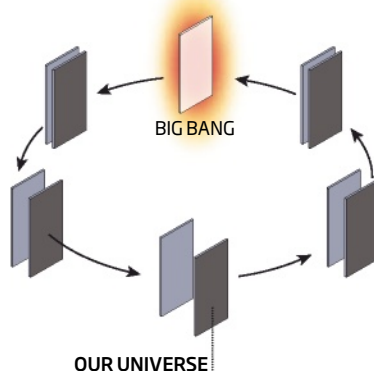


Three ways to an eternal universe

There are several ways which the universe could have existed forever. Each of these arguments has its flaws

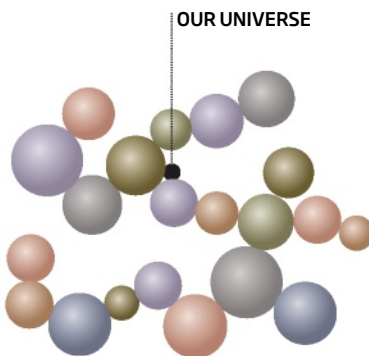
CYCLIC UNIVERSE

4D universes repeatedly collide together in a fifth dimension creating big bang events



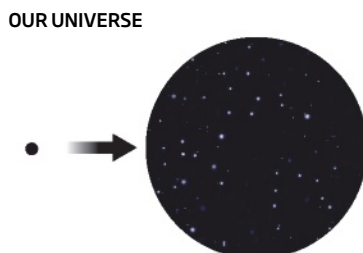
ETERNAL INFLATION

Universes have always been inflating from the vacuum and always will



EMERGENT UNIVERSE

Tiny universe has existed forever but blew up to its present size



this quantum instability result “is not unreasonable”. However, he is not sure this inevitably implies a beginning to the universe. “It strikes me that this is a much deeper question, requiring at least a mature quantum gravity theory,” he says. “Unfortunately, we do not have that.”

In the context of known physics, however, Vilenkin and Mithani conclude that, whatever way you look at it, the universe cannot have existed forever so must have had a beginning. But how did it begin? According to Vilenkin, quantum theory has a solution because it permits something to pop out of nothing – with that something being a small universe that starts to inflate, cycle or hang for an extremely long time before inflating.

Can we really be sure now that the universe had a beginning? Or are we in for an infinite cycle of belief and disbelief over the matter? “For the first time in history, we have the tools to address the origin question scientifically,” says Vilenkin. “So I have a feeling we are getting near to the truth.”

Any hope of us observing the ultimate origin is fading, however. Soon after Vilenkin and Mithani published their argument, physicist Leonard Susskind of Stanford University in California responded with two papers. In them, he says that a beginning, if it did indeed occur, is likely to have been so far in the past that for all practical purposes the universe has been around forever.

He argues that because space inflates exponentially, the volume of the vacuum at later times is overwhelmingly greater than at earlier times. With many more bubble universes in existence, chances are that the patch of vacuum we call home formed later on too. The true beginning is likely to have been an awfully long time ago – so far away, that no imprint on the universe has survived. “I find it a paradoxical situation to say that there must have been a beginning, but it is with certainty before any nameable time,” says Susskind.

Vilenkin acknowledges this. “It’s ironic,” he says. “The universe may have a beginning but we may never be able to know exactly what the beginning was like.”

Still, cosmologists have plenty of other big questions to keep them busy. If the universe owes its origins to quantum theory, then quantum theory must have existed before the universe. So the next question is surely: where did the laws of quantum theory come from? “We do not know,” admits Vilenkin. “I consider that an entirely different question.” When it comes to the beginning of the universe, in many ways we’re still at the beginning. ■

triggered to inflate. Such scenarios do arise in string theory, so the idea isn’t totally out of the blue. “It’s a somewhat desperate scenario,” says Vilenkin.

To model an eternally slumbering emergent universe is not straightforward. In the same way that Einstein’s static universe was unstable and needed the extra ingredient of cosmic repulsion, Ellis and Maartens can only stabilise theirs with two weird ingredients: a vacuum with negative energy, and fault lines in space-time known as domain walls that are a feature of some models of particle physics. Domain walls should leave an imprint on the temperature of the cosmic microwave background radiation, which has not been seen, but this might be explained if they were diluted away by inflation.

Spontaneous collapse

Vilenkin and Mithani are critical of Ellis and Maartens’s approach. “At first sight it appears they have concocted a stable universe,” says Vilenkin. “However, we find that it’s only stable if you ignore the effects of quantum theory.”

According to quantum theory, the universe cannot stay at its minimum size forever – there is a chance it would spontaneously collapse. “Although the probability may be very small, since an infinite amount of time is available, it is inevitable,” says Vilenkin. “Consequently, if we live in an emergent universe, it cannot have existed forever.” According to Maartens,

Bang goes the theory

Just what kind of a bang was the big bang?
Amanda Geffer discovers trouble with the start of time

YOU know how sometimes you meet somebody and they're really nice, so you invite them over to your house and you keep talking with them and they keep telling you more and more cool stuff? But then at some point you're like, maybe we should we call it a day, but they just won't leave and they keep talking and as more stuff comes up it becomes more and more disturbing and you're like, just stop already? That's kind of what happened with inflation."

Max Tegmark's voice is animated as he talks about the idea underpinning our story of the universe's origins. The theory of cosmic inflation states that in the first fraction of a second after the big bang, the universe's fabric expanded faster than light. Without a trick like that, we have difficulty squaring some crucial features of today's cosmos with a universe that began as a hot, dense soup and has been growing and cooling since.

Yet Tegmark, a cosmologist at the Massachusetts Institute of Technology, is not the only person asking whether inflation has outstayed its welcome. The theory has received a shot in the arm with the tentative discovery of gravitational wave imprints in the cosmic microwave background (CMB) radiation, the afterglow of the big bang (see "Testing, testing", page 15). But for all its attractions, the theory has unpalatable consequences – so much so that they threaten to undermine our entire understanding of the cosmos. Debate has been reopened on a question many thought had been settled: what kind of a bang was the big bang?

When inflation was proposed in 1980 by Alan Guth, then a young postdoc at Cornell University in Ithaca, New York, it was a godsend. Studies of the CMB were causing serious headaches among proponents of the big bang.

This radiation seems in itself to be a spectacular confirmation of the big bang theory. Some 380,000 years after the universe's origin in a "singularity" of

unimaginable density and temperature, the cosmos had expanded and cooled enough for the first atoms to form. Photons of light pinged off these atoms and have been travelling in all directions through the cosmos ever since. Sitting in our particular spot some 14 billion years later, we see them as a background radiation that suffuses the sky at an almost uniform temperature of 2.7 kelvin.

But that success hides some problematic details. One is that you can measure the cosmic background 10 billion light years away in one direction and 10 billion light years in the other and still observe that pleasing uniformity. Run a simple story of a steadily expanding cosmos backwards, however, and it would take 20 billion years for these patches of space to meet – more than the age of the universe. In a simple big-bang universe, they were never close enough to equalise their temperatures, and that uniformity is a highly improbable coincidence.

Then there is the universe's geometry. Information encoded in the cosmic background shows space is extremely "flat": Euclidean geometry reigns supreme and parallel lines never meet. This flatness is also highly unlikely, given what we know about gravity and its compulsion to warp space. Again, a simple big bang cannot explain it.

Inflationary pressure

An inflationary big bang can. At the beginning of time, so the idea goes, all that existed was a quantum field called the inflaton. It found itself in a "false vacuum" – a state that is temporarily stable, but not the lowest, true-vacuum energy state. It is as if the inflaton were poised on a small plateau on a steep mountainside. All things being equal, it could rest there undisturbed, but the slightest jiggle would send it careering down towards the true vacuum below.

In the universe's first fraction of a second, all things were not equal. Random quantum ➤





Big bang 1: Inflationary cosmology

By rapidly pushing apart the early universe, a period of inflation can explain why distant parts of the cosmic microwave background look like they came from the same place

fluctuations in energy provided just the jiggle to set the inflaton on its way. As it fell towards the true vacuum, it generated a kind of repulsive gravity that pushed the space outaround it. The further it fell, the more it pushed until space was ballooning outward at a speed far greater than that of light.

This is physically all above board. Einstein's relativity forbids objects from travelling faster than light through space, but places no constraints on what space itself can do. And when the inflaton hit rock bottom, all the kinetic energy it acquired in its headlong descent poured into the universe, creating the matter and radiation that went on to form stars, planets and, eventually, us. All this happened in considerably less than the blink of an eye: in just 10^{-33} seconds, the observable universe ballooned over 20 orders of magnitude in size, from a diameter about a billionth that of an atomic nucleus to a mini-cosmos about a centimetre across (see diagram, right).

Unwanted universes

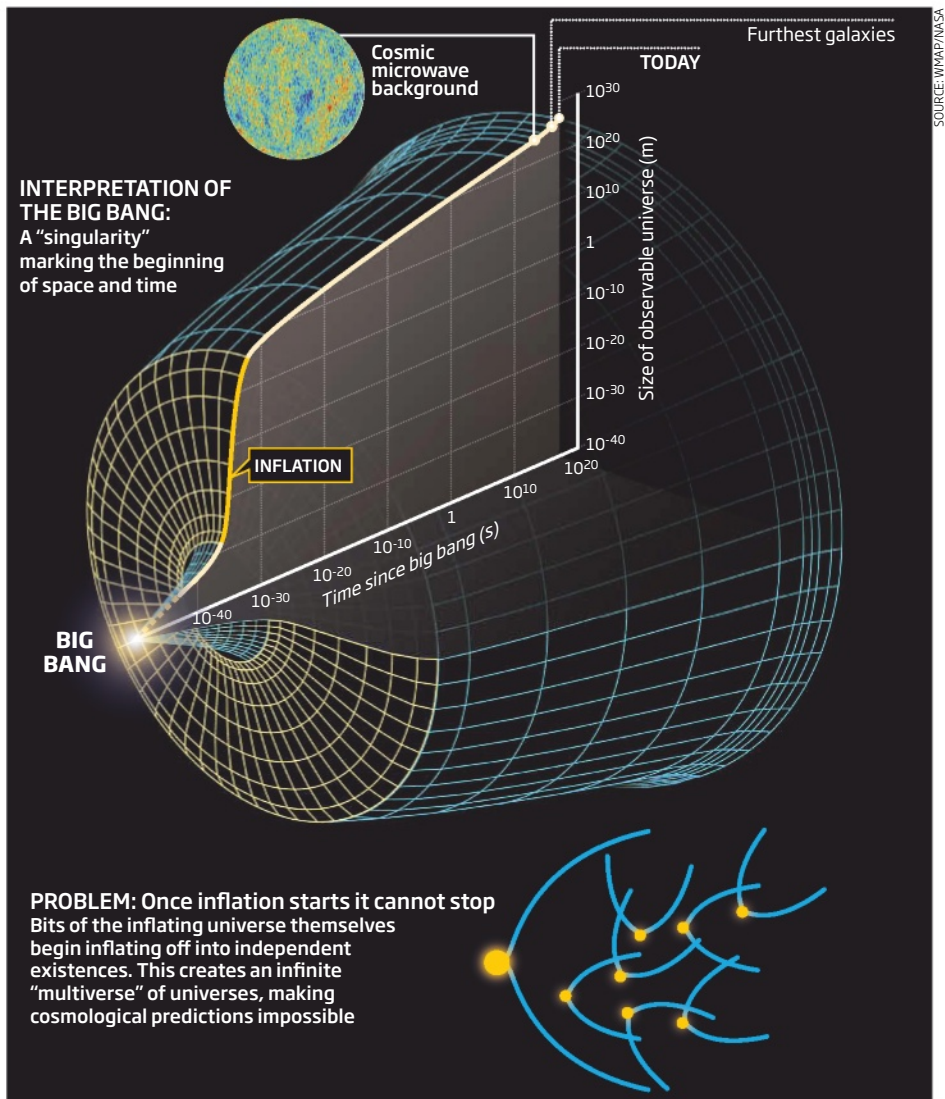
In one fell swoop, inflation solved the big bang's problems. Those patches of sky no longer need a 20-billion-year rewind to have met and mingled: inflation gave them the shove-off they needed to ensure they arrived far faster at the far reaches of the cosmos. And that absurdly unlikely flatness is nothing of the sort: inflation makes the universe so large that any measurable region must look flat, the same way the ground at your feet looks flat even though Earth's surface is curved.

Inflation's munificence didn't stop there. By inflating tiny quantum fluctuations in the density of the cosmos to astronomical proportions, it produced a blueprint for how stuff clumped into ever-larger agglomerations of matter such as the galaxies we see today.

"That would have been the perfect point for inflation to bow, wait for applause and exit stage left," says Tegmark. But that didn't happen. Instead, inflation kept on predicting still more things – things that nobody wanted.

Things like other universes.

The problem is that once inflation starts, it is nearly impossible to stop. Even in the tiny



pre-inflation cosmos, quantum fluctuations ensured that the inflaton field had different energies in different places – a bit like a mountain having many balls balanced precariously at different heights. As each one starts rolling, it kicks off the inflation of a different region of space, which races away from the others at speeds above that of light. Because no influence may travel faster than light, these mini-universes become completely detached from one another. As the inflaton continues its headlong descent in each one, more and more bits of space begin to bud off to independent existences: an infinite "multiverse" of

universes is formed (see diagram, above).

This is not good news for our hopes for cosmic enlightenment. In a single universe, an underlying theory of physics might offer a prediction for how flat the universe should be, say, or for the value of dark energy, the mysterious entity that seems to be driving an accelerated expansion of the universe. Astronomers could then go out and test that prediction against observations.

That's not possible in an infinite multiverse: there are no definite predictions, only probabilities. Every conceivable value of dark energy or anything else will exist an infinite

number of times among the infinite number of universes, and any universal theory of physics valid throughout the multiverse must reproduce all those values. That makes the odds of observing any particular value infinity divided by infinity: a nonsense that mathematicians call “undefined”.

At first, cosmologists hoped to make sense of these infinities by taking a finite snapshot of the multiverse at some particular time, and then extrapolating the relative probabilities of various observations out to later and later times and an ever-larger number of universes. Einstein stymied that approach. His relativity means there is no single clock ticking away the seconds of the multiverse, and there are an infinite number of ways to take snapshots of it, each giving a different set of probabilities. This “measure problem” destroys inflation’s ability to make predictions about anything at all, including the smoothness of the cosmic background, the curvature of space, or anything else that made us believe in the theory in the first place.

“We thought that inflation predicted a smooth, flat universe,” says Paul Steinhardt of Princeton University, a pioneer of inflation who has become a vocal detractor. “Instead, it predicts every possibility an infinite number of times. We’re back to square one.” Tegmark agrees: “Inflation has destroyed itself. It logically self-destructed.”

Sean Carroll, a cosmologist at the California Institute of Technology in Pasadena, is more circumspect. “Inflation is still the dominant paradigm,” he says. “But we’ve become a lot less convinced that it’s obviously true.” That’s not just because of the measure problem, he says. More basically, we don’t know what an inflaton field is, why it was in a false vacuum and where it and its energy came from. Having an inflaton field perched so perfectly and precariously atop that mountainside seems no more likely than the flukes the idea was intended to explain. “If you pick a universe out of a hat, it’s not going to be one that starts with inflation,” says Carroll.

So if it isn’t inflation that made the universe as we see it today, what did? Are there other sorts of big bang whose predictions square

with our observations, but don’t have inflation’s unwanted side effects?

Perhaps. In 2001, Steinhardt was one of the first to suggest an alternative, together with his colleagues Justin Khoury, Burt Ovrut and Neil Turok. Their idea was to revisit our interpretation of the big bang. Rather than marking a singularity at the absolute beginning of space and time, it was just a recent event in a much longer history. The inspiration for this idea came from string theory, the most widespread approach to get Einstein’s general theory of relativity, which best describes space and time, to play nicely with quantum mechanics, which best describes everything else.

String theory proposes that the various particles that make up matter and transmit forces are vibrations of tiny quantum-mechanical strings, including one that produces a “graviton”, an as-yet-undetected particle that transmits gravity. It also predicts the existence of extra dimensions beyond the four of space and time we see.

Brane drain

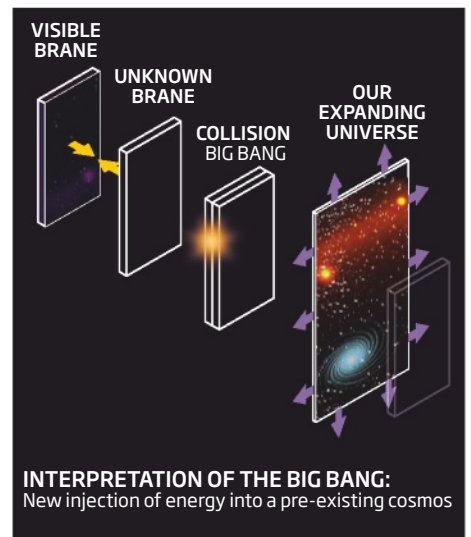
This leads to the possibility that our 4D cosmos is situated on a “brane”, a lower-dimensional object floating in a higher-dimensional space. In this picture, ours is not necessarily the only brane: two branes floating a microscopic distance away from each other can form the boundaries to a five-dimensional space in between, like two slices of bread bounding a sandwich. Steinhardt and his colleagues’ idea was that every few trillion years or so neighbouring branes float together and touch, resulting in an explosive exchange of energy as the fifth dimension briefly disappears into a singularity, reappearing as the branes move apart again (see diagram, right). Our 4D brane receives a huge injection of energy in this collision, something we interpret as a big bang. It doesn’t represent the absolute beginning of our space and time, however, but a reinvigoration of a cosmos with an eternal existence.

This “cyclic model” of the big bang does much of what the inflationary big bang ➤

“Inflation ends up predicting every possible universe an infinite number of times. We’re back to square one”

Big bang 2: Colliding worlds

If our universe is a 4D “brane” floating in a higher-dimensional space, it might collide with other branes, acquiring energy that we interpret as the big bang



"We use a half-assed version of quantum mechanics when we do cosmology – maybe that isn't good enough"

was invented to do. "You can solve problems that would have been intractable without inflation given only 14 billion years," says Steinhardt. The branes are essentially flat to begin with, so the flatness problem disappears. When they clash, they hit at almost the same time everywhere, so the energy that will form matter and radiation pours in almost uniformly, creating a nearly homogeneous cosmos. Small quantum energy fluctuations are all that are needed to give enough density variation to eventually seed galaxies. And because the idea needs no multiverse, the measure problem disappears.

If nothing else, the cyclic model introduced some competition into the big bang market. "It shows that you're not stuck with inflation – other ideas are possible," says Steinhardt. "But whether or not you like this particular alternative is a matter of taste."

Not everyone did. Models of the big bang that involve a singularity in our space-time, including the inflationary big bang, neatly excuse us from explaining what happened at the universe's beginning: the singularity is a place where the universe falls off the cliff of existence and the laws of physics break down. But in the cyclic model, we must explain how the fifth dimension survives its momentary

lapse into a singularity. "To me, it doesn't seem to work," says Thomas Hertog of the Catholic University of Leuven (KUL) in Belgium, who worked on the idea for a couple of years. "The calculations suggest that the transition through the singularity is very unlikely."

The many clashes between branes that the model implies just compound the problem, says Carroll. "If you follow the cyclic universe backward in time, the conditions that you need become more and more special, or unlikely." The way to solve that problem is to postulate some kind of beginning that provided a special set of conditions – but that seems to defeat the object of the theory.

Many histories

So if neither inflation nor the cyclic universe can deliver a plausible beginning to the cosmos, what can? Carroll suggests a middle way. Steinhardt might have had the right idea in seeking an answer to the big-bang conundrum in the unification of general relativity and quantum mechanics, he says. But the answer could lie in using those ideas not to replace inflation, but to make it better.

"We use a half-assed version of quantum mechanics when we do cosmology," Carroll says. Inflation is a theory about space-time and gravity, so it is anchored in general relativity. It incorporates a few aspects of quantum physics, such as the uncertainty fluctuations that push the inflaton off the mountain ledge, but as we lack a sure-fire way of connecting relativity and quantum theory, it remains a "semi-classical" theory. "Maybe that's not good enough," says Carroll.

But how can we put the quantum into cosmology? Just after inflation had burst on to the scene, physicists Stephen Hawking of the University of Cambridge and James Hartle of the University of California, Santa Barbara, made a stab at it. In quantum physics, when a particle travels from A to B it doesn't take a single path but can pass along two or more paths simultaneously, interfering with itself at the other end as if it were a wave. To find out which path we are most likely to observe, we must add together the quantum-mechanical

"wave functions" encoding each possible path, working out how their individual peaks and troughs cancel and amplify each other. Encoded within this total wave function is everything we need to know about the quantum particle at B, including the probabilities for the outcomes of any measurement we choose to make.

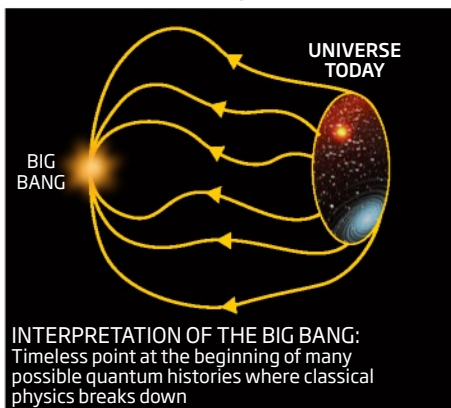
Hawking and Hartle argued that a similar approach could be applied to the universe as a whole. Point B is the universe we see today. Looking back towards its origin, we can trace many valid histories of its expansion back towards a point – point A – where semi-classical physics breaks down and quantum space and time become so gnarled that the two cease to be clearly distinguished.

This point is no longer a beginning of time in a singularity as in the standard inflationary cosmology, but a timeless point where the universe – or rather a superposition of all possible historical universes – pops into existence from nothing with all its laws of physics intact. Because of this lack of a clear beginning, Hawking and Hartle called it the no-boundary proposal.

Following the rules of quantum mechanics, they added up all the possible histories that began in a universe with no boundary and ended in the universe we see today. The idea is a kind of multiverse in reverse: a single universe with multiple histories (see diagram, left). The resulting wave function gets rid of

Big bang 3: Uncertain history

By adding the possible quantum histories of the universe together, we can work back from today's cosmos to the universe's origin





proposal was true. It was difficult to see where it fitted in to the sort of unifying theoretical constructs, such as string theory, which are needed to explain events in the early, high-energy days of the universe.

That might have changed, thanks to one of the most profound ideas to come out of string theory in recent years: the holographic principle. This states that the physics of a 4D universe such as ours, including gravity, is mathematically equivalent to the physics on its 3D boundary without gravity. The implication is that the world we see around us is nothing but a holographic projection of information from the edge of reality. It sounds implausible, but the principle pops up not just in string theory, but in almost any approach to unifying relativity and quantum theory dreamed up so far.

Although the no-boundary proposal says that the universe has no boundary in the far past, it does give a boundary in the infinitely far future. By calculating the physics on this boundary, Hertog extracted the probabilities of all the possible universes that can emerge as its holographic projections. Remarkably, the probabilities for things like the homogeneity of the cosmic background or the amount of dark energy are the same as those that you get from the no-boundary wave function. This supplies a direct connection between string theory, the most popular route towards a theory of everything, and the no-boundary proposal, which produces inflation naturally.

"Originally the no-boundary wave function was sort of picked out of thin air," says Hertog. "But now we see that it lies at the heart of the holographic principle. This is very encouraging for inflation." Together with Hawking and Hartle, he set out the thinking in a paper.

Cosmologists are still digesting the proposal, with some questioning whether the assumptions it makes are justified. Guth says that as yet no one is sure about the validity of the specific holographic correspondence Hertog, Hawking and Hartle have employed. "It's certainly a worthwhile line of research, but what they are trying to establish is a very difficult thing," he says.

We are not yet there, at the true story of the beginning of the universe. But it seems undeniable now that working out whether inflation is the imposter doubters such as Tegmark believe it is will depend on finding some way to consistently apply quantum theory to the fabric of the universe. Only then will we truly know what kind of a bang the big bang was. ■

the measure problem, as it encodes a unique set of probabilities for anything we might observe. And because the admissibility of the histories is determined by what we see in the cosmos today, problems such as the flatness of space-time or the homogeneity of the cosmic microwave background cease to be problems: instead, they are the inputs to the theory.

The no-boundary universe had its attractions for many physicists. "If it turns out the universe has to have an origin, I find this initial state to be a very plausible one," says

Guth. And as far as Guth's theory of inflation is concerned, the no-boundary proposal turned up a pleasant surprise. As Hartle, Hawking and Hertog showed in 2008, although the theory was not dreamed up with inflation in mind, it crops up naturally along many paths the universe could have taken to get here. "You can calculate the probability that inflation occurred, and it turns out that probability is very high," says Hertog.

That all sounds very neat, but there was still no reason to believe the no-boundary

TESTING, TESTING

Theorising about the big bang can only take us so far: we need observations and experiments to determine the true story of the universe's beginnings. Fortunately, a number are in progress that might help us discriminate between rival theories (see main story).

If the extraordinarily powerful burst of early expansion implied by the theory of cosmic inflation really did occur after the big bang, it would have sent ripples through the fabric of space-time. These gravitational waves would have had subtle effects on the

polarisation of the microwave radiation that makes up the background - and we have instruments that should be capable of spotting them. First out of the blocks was BICEP2, a telescope based at the South Pole. In March 2014 researchers working on BICEP2 data reported results consistent with a direct sighting of gravitational waves. By the time you are reading this, the European Space Agency's Planck satellite, launched in May 2009 to make the most detailed measurements of the cosmic microwave background yet, may

have confirmed the sighting. Meanwhile, Celine Boehm of Durham University, UK, and her colleagues have proposed that the high-energy collisions at the Large Hadron Collider at CERN near Geneva, Switzerland, could be used to test whether the inflaton field and its associated particle, the inflaton, existed and if so what it looked like.

The particle itself is far too massive to be made in the collider, but its existence might influence the decay of other, less weighty, particles - and that could be tested.

Inflation handily explains how the universe expanded after the big bang. Is it all a bit too convenient? **Maggie McKee** investigates

Pop-up universe

IT WAS still dark and bitterly cold when Paul Steinhardt met Anna Ijjas at her office in Harvard on 21 March 2013. They had risen before dawn to watch researchers in Paris announce the first major results from the European Space Agency's Planck satellite, which has been mapping light from the universe's infancy since 2009.

Other missions have studied this light, aka the cosmic microwave background. But cosmologists the world over were itching to find out what Planck had seen because its detectors are at least three times as sharp as those on any previous satellite. It therefore provides the best-ever portrait of the universe not long after the big bang.

Would Planck lend weight to our leading theory of a well-mannered universe that underwent a precocious growth spurt known as inflation? Or would it force us back to the drawing board to understand our cosmic past?

"The overall conclusion is that it is an extremely good match to the Planck data," announced George Efstathiou at the Paris meeting. "If I were an inflationary theorist I'd be very happy."

But after reading the Planck team's paper, which went online shortly after the briefings, Ijjas reached the opposite conclusion. And

soon she convinced Steinhardt and Abraham Loeb, the head of Harvard University's astronomy department, that the results were troubling. "Planck has created some problems for inflation that weren't there before," says Steinhardt, a cosmologist at Princeton University.

Others disagree with that conclusion. But if Ijjas, Loeb and Steinhardt are right and inflation really is in difficulties, it would spell disaster for our entire understanding of the cosmos. Either way, Planck is driving us to take a harder look at the theory.

Smooth operator

Inflation proposes that our corner of space expanded by a factor of 10^{25} or more in a tiny fraction of a second soon after the big bang. To get a sense of how incredible this expansion is, imagine the full stop at the end of this sentence growing to 1 million light years across in the blink of an eye. Yet this remarkable theory has become part of our standard picture of cosmology because it seems to elegantly account for a number of astronomical puzzles.

One is the astonishing evenness of the universe. Observations show that the

microwave background radiation filling the cosmos is the same temperature everywhere to within 0.0003 °C. That's surprising because it is also true for parts of the sky that appear too far away for light to have had time to travel between them. So there is no way heat radiation could have evened out the roiling temperatures present at the universe's birth. Inflation provides an explanation – these parts of space had been in contact before inflation began, and had simply been pushed to great distances when the universe suddenly swelled up.

And yet, if the universe was identical everywhere then there would be no structures like galaxies, stars and planets. Inflation has an answer for this too. Quantum fluctuations in energy continually produced small dents in space-time. When inflation began soon after the big bang, it stretched these dents out to cosmic proportions, and they acted as nests for matter to fall into and eventually grow into the structures we see today. These nests appear as spots that are slightly warmer and cooler than the average temperature of the cosmic microwave background (CMB). They were first measured by the COBE satellite in 1992 and later by Planck in unprecedented detail. "Planck confirmed the quantum





origin of the universe's structure," says Slava Mukhanov of Ludwig Maximilians University Munich in Germany.

So what's not to like about the theory? Well, for one thing, it's not clear what actually does the inflating. One view is that it's caused by a quantum field associated with a hypothetical particle called the inflaton. At each point in space, the idea goes, the inflaton field has a certain strength, much as an electric field varies with distance from an electric charge.

The energy that an electric field stores is

agrees: "If you require the universe to be homogeneous to start with, you're not solving that problem."

This adds up to what Ijjas, Loeb and Steinhardt call an "initial conditions problem" for plateau models of inflation. They also suffer from being unlikely, the trio say, in the sense that models that produce more inflation and produce it more easily are more likely to have occurred. Plateau models fail on that account too. "The models which are now favoured by the data are theoretically

exponentially less likely," Ijjas says.

Martin Bucher of Denis Diderot University in Paris, who was one of the leaders of the Planck team that focused on inflation,

agrees that "not every potential is plausible". But he says the measurements do not rule out all of the models with steeper curves. That's because inflation is thought to have pushed most of the universe so far away from us that light from those distant regions has not had time to reach us in the 14 billion years or so since the big bang. The hot and cold spots produced in the very early phases of inflation would be larger than our observable universe, giving us a view of only part of it. "There's tons and tons of inflation that you don't see," says Bucher.

An infinity of universes

The energy curve derived from our observable universe seems to look like an upturned shallow soup bowl. But it could have a steep part too, starting off steeply and then flattening off before falling again – looking for all the world like a Hershey's Kiss chocolate.

Andrei Linde of Stanford University discussed a model with such an energy curve at a physics conference in Santa Barbara, California, in April 2013. Because the potential energy starts out high, there is no need to assume the newborn universe was the same on very large scales. "You have no problems with initial conditions," he pointed out.

According to Linde, who is one of the architects of inflation theory, creating the "shoulder" between the steep and flat part of the curve means tuning the potential by about 1 per cent to fit the Planck measurements. Yet Ijjas and her colleagues say such tweaks are undesirable. "You can always go back and find a potential that will give you what you need, but if you're just reverse-engineering the initial conditions, that's not a good approach

"Once inflation starts, it is impossible to turn off completely, leading to an infinite number of universes"

proportional to the square of its strength, a relationship that in a graph looks like a U-shaped parabola. It means that where the field is twice as strong, it harbours four times the energy. Yet the inflaton's behaviour is up for grabs. If it exists, the value of its potential energy dictated how fast space expanded due to inflation. So it must have been positive when inflation started and then dropped to zero when it ended. But how high did the energy reach, and how did it vary from inflation's start to its end? Hundreds of models have suggested different answers to these questions. "You're allowed to invent any energy curve you want," says Steinhardt.

Previous, lower-resolution maps of the CMB weren't detailed enough to give a clear picture of how the process might have played out. If you imagine the speed of inflation changing in the same way as a ball rolling down a hill, it wasn't clear whether the hill was like the inside of a bowl – amounting to inflation that is fast to begin with and slows down at the end – or the outside of an upturned one. Also unclear was whether it was a shallow soup bowl or a tall vase.

Planck's improved view of the distribution of CMB temperatures points to a slide down the outside of a shallow bowl. That makes things difficult, says Ijjas, because such "plateau models" are complicated. They need to be finely tuned to provide sufficient inflation and to avoid a scenario in which the universe collapses into a black hole. Some plateau models therefore call for the universe to have started out a billion times smoother than do simpler models with a steeper energy curve. "That's a very big problem, because inflation was introduced to explain the smoothness of the universe," says Ijjas. Loeb

to doing basic science," says Loeb. "It doesn't lead you to believe that the model is predictive."

Their third criticism of inflation concerns what Planck doesn't see. Once inflation starts, it is thought to be impossible to turn off completely, leading to an infinite number of universes. But the Planck data shows no obvious sign of them. A giant cold spot has been seen, and some researchers claim it might be a sign of an interaction with another universe. However, its presence may be nothing more than a statistical fluke. Since inflation is supposed to make universes





it is affected by the Higgs field, is not in its lowest possible energy state. Instead, it appears to be metastable, meaning that at some point it could decay into its lowest energy state. If that happens, it would cause the universe to collapse into a point of infinite density.

Cosmic phoenix

The diagnosis of metastability for the Higgs vacuum is itself on shaky ground, but if it is metastable that presents two problems for inflation. Add the Higgs and inflaton fields, for example, and you wouldn't necessarily have the energy needed to inflate in the first place, says Steinhardt. And if inflation did take place and was extremely energetic, it could have caused the Higgs vacuum to decay and so wiped out the universe.

So what to do? Although Steinhardt was one of the founders of inflationary theory, he has an alternative idea that avoids these problems. He and Neil Turok of the Perimeter Institute for Theoretical Physics in Waterloo, Ontario, Canada, have worked on a theory in which the universe alternates between periods of

"A reincarnated cosmos does the same things over the course of its past life that inflation does in under a second"

expansion and contraction. In this cyclic model, the big bang is not a birth but a rebirth, a cosmic phoenix rising from the ashes.

The reincarnated universe accomplishes the same things over the course of its past life that inflation does in a fraction of a second. Its long, slow contraction, for example, brings together far-flung regions of space, explaining why they appear eerily clone-like today. "There's so much more time for physical processes to occur," says Steinhardt. In a subsequent version of the theory, developed with Itzhak Bars at the University of Southern California in Los Angeles, the contraction happens so slowly and gently that the Higgs vacuum can remain in its metastable state from one cycle to the next.

Other theorists point out that the cyclic model has a number of shortcomings, which Steinhardt and his colleagues are working to address. "I am not wedded to inflation; in fact, I think we are in need of a grander idea," says Michael Turner of the University of Chicago. "I am not convinced the cyclic model is that grander idea."

Certainly for some physicists, inflation is at

an impasse. The bad news is that further experimental tests may not be definitive. Inflation should have shaken space-time itself, triggering gravitational waves that spread throughout the universe. But just how violent this shaking was, and therefore how likely we are to be able to detect the waves today, varies according to each model. Planck and other missions have been looking for the subtle imprint these waves left on the CMB. Early results from the BICEP2 telescope at the South Pole suggest that gravitational waves are indeed imprinted on the CMB.

Such waves are not compatible with the cyclic model, since the contraction should have been too gentle to rattle space-time significantly: a confirmed detection of gravitational waves will rule out the cyclic model.

But even if no signal is confirmed, inflation still has a fighting chance. It might just mean that inflation was less energetic than generally thought, says Sabino Matarrese, a Planck team leader at the University of Padua in Italy.

The fact that inflation contains a multitude of predictions works against it, because that makes it difficult to falsify. "If you find something that has in it the potential to explain everything, then that is problematic," says Benjamin Wandelt at

Sorbonne University in Paris and a Planck team leader. "Unsinkable theories are highly suspect in science."

The problem, he and others say, is that there is no well-agreed guiding principle from particle physics illuminating what the inflaton should be. "The truth is, we don't have a compelling model for inflation or even a class of compelling models," says Turner.

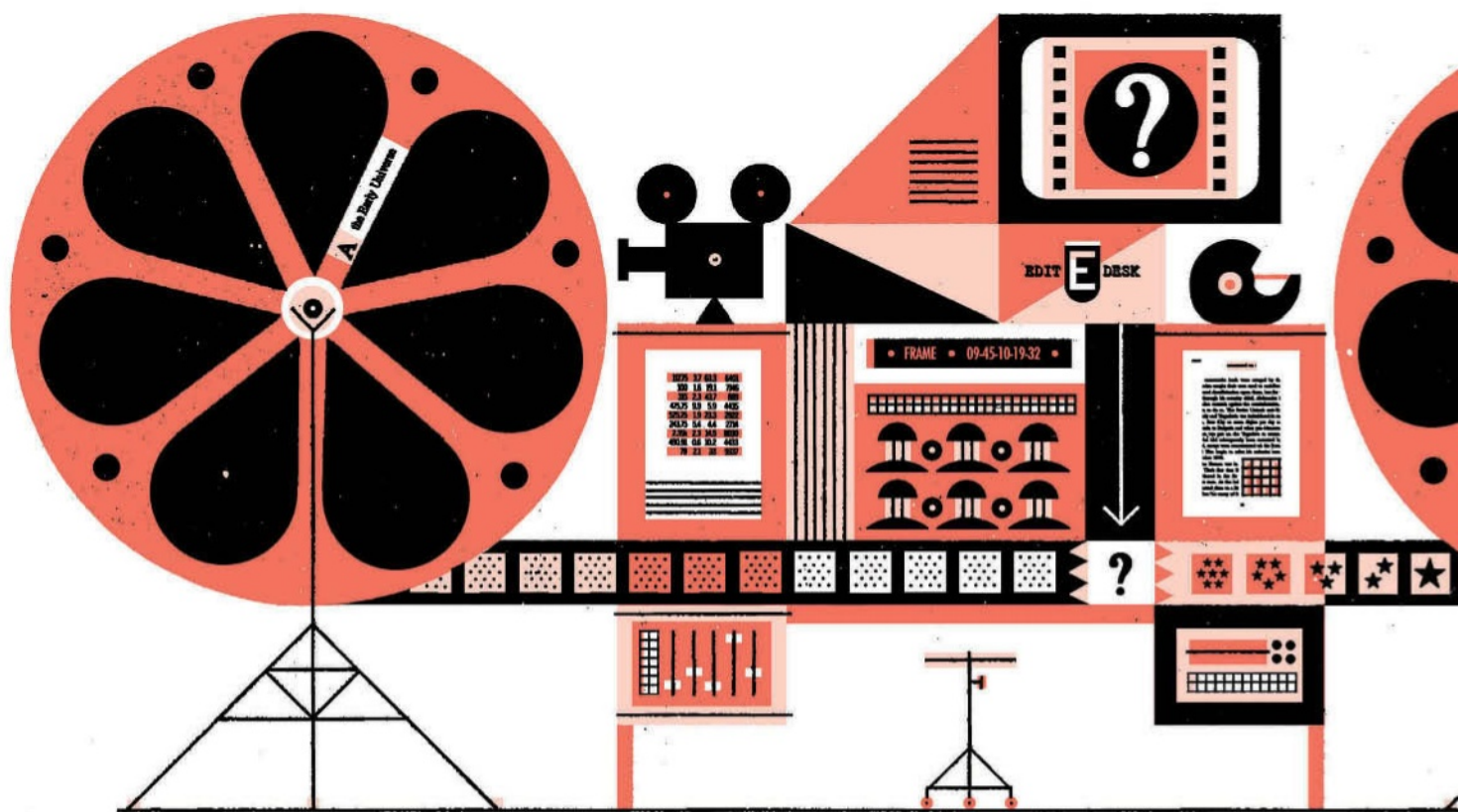
Matarrese thinks some progress could be made if physicists pin down why the universe is made up of more matter than antimatter, since the energy in the inflaton field should have been converted into particles of both when inflation ended. If gravitational waves are not confirmed at the expected energy scale after a careful search, "inflation would be in trouble," says Matarrese.

The uncertainty over inflation is "just the tension that we live with at the moment", says Wandelt. Passions can run high in the debate over its existence because missions like Planck are providing a window into the universe's first moments and researchers are anxious to get a better view. "It is a very, very special time in the history of human ideas," he says. ■

reproduce like rabbits, Steinhardt argues, why would there be only one cold spot?

A final problem highlighted by Ijjas and her colleagues relates to the discovery of the Higgs boson in 2012. Most of the energy in the newborn universe is considered to have come from the quantum field associated with the inflaton. The field associated with the Higgs had been thought to have little effect. But that picture might have to change, says Steinhardt, in light of the particle's mass.

Curiously, its value of 126 gigaelectronvolts suggests that the vacuum state of today's universe, known as the Higgs vacuum because



The missing reel

The universe's story is a true epic – but up till now we have been lacking some decisive early scenes, say astrophysicists **Abraham Loeb** and **Jonathan Pritchard**

ASTRONOMERS have a great advantage over archaeologists: they can see the past. The finite speed of light means that the further away a cosmic object is, the longer its light takes to reach us. An image recorded by our telescopes today tells us how that object looked long ago when its light was emitted.

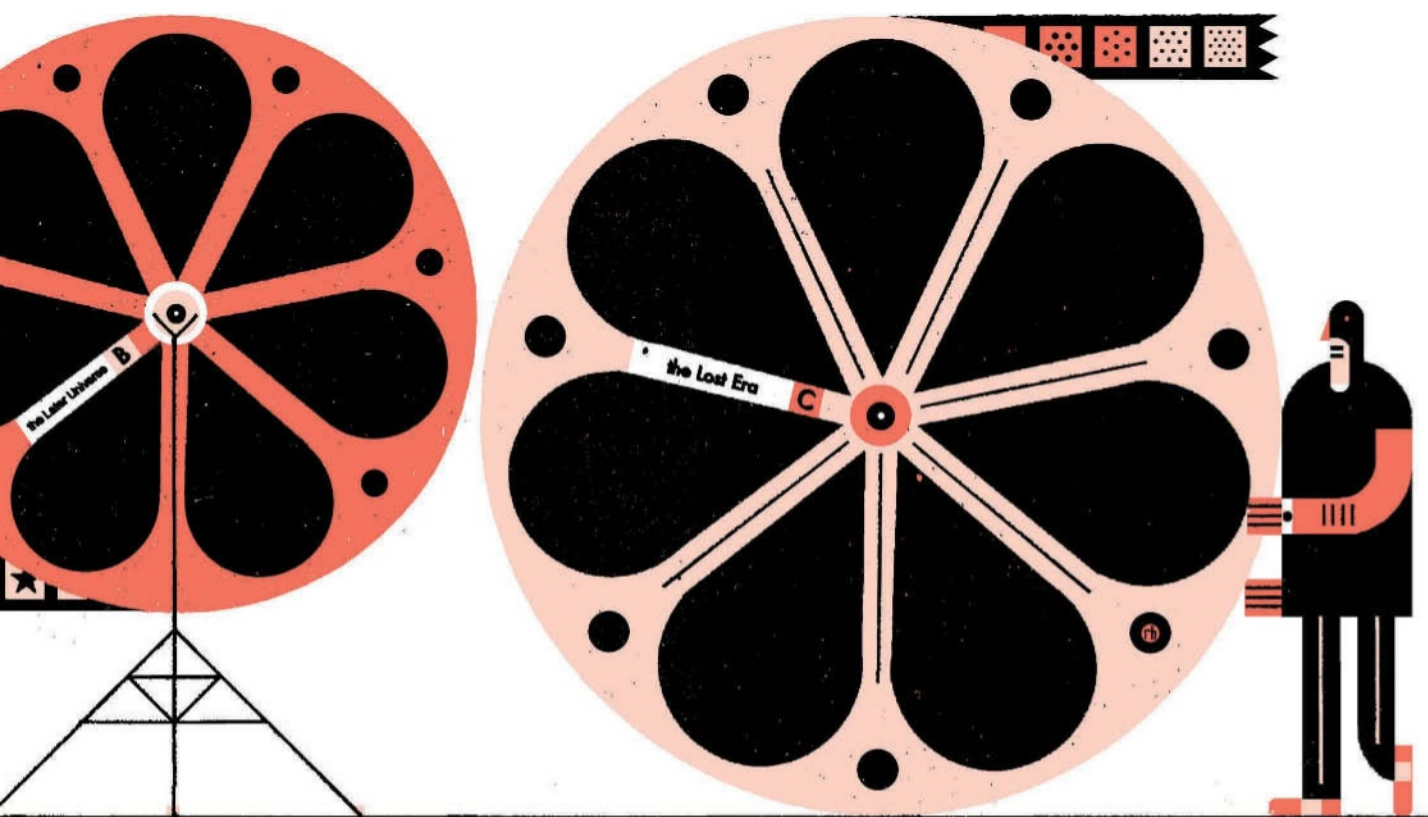
Using our most powerful telescopes on Earth and in space, we can now trace the universe's story back to a time when it was just 500 million years into its 13.8-billion-year life. Meanwhile, a single flash that bathed the cosmos in light 400,000 years after the big bang provides us with an isolated exposure of

the universe in this infancy. This infant universe is, like a newborn baby, almost featureless, yet to assume the characteristics that will mark it out in later life. When our telescopic cameras pick up its story again, however, it is recognisably its adult self – stars, galaxies and clusters of galaxies already populate its reaches.

What happened in between, during the universe's turbulent, formative adolescent years? That has long been a matter of conjecture. Now, thanks to a combination of new instruments and refined observational techniques, the missing reel of the universe's

story is about to be slotted in – presented in the crackle and hiss of giant radio waves.

We already have a rough storyboard for this missing adolescence. Clues are encoded in the cosmic microwave background, the radiation suddenly liberated 400,000 years into the universe's life. At this "epoch of recombination" the cosmos had cooled enough for protons and electrons to form neutral hydrogen, which scattered light in all directions. Tiny variations in this radiation's temperature show that the atoms were spread not uniformly, but in almost imperceptible clumps. Gravity's pull, our story goes, caused these lumps to consolidate and



grow and eventually ignite into stars. These stars, in turn, felt a mutual attraction, slowly forming ever-larger galaxies over the course of a few hundred million years.

As they did so, the universe underwent a final change in its character: high-energy radiation broke down hydrogen formed at the epoch of recombination, freeing up electrons and protons. This “epoch of reionisation”, which is thought to have ended some 700 million years after the big bang, marked the coming of age of the cosmos we see today.

Convincing as this plot development is, in the absence of observational evidence it is a tale too confidently told. Many significant details of the universe’s evolution remain sketchy – and in some cases wholly obscure.

For a start, what made the first galaxies? The first stars, one might reasonably think. But the first stars must have been oddballs. Unlike every subsequent generation, they grew up in a pristine environment consisting solely of hydrogen and helium, the only elements the big bang forged in large quantities. Nuclear reactions within these pioneering stars created the heavier elements, such as carbon, oxygen and silicon, that went into the mix for later stars such as our sun –

and, ultimately, our planet and ourselves.

From what we can tell, these first stars were bloated monsters up to 100 times the size of our sun that lived fast, burned bright and died within a few million years. Was that long enough to form their own galaxies, or to influence their formation – or was it their less flamboyant successors that first found safety and stability in numbers?

Cosmic monsters that have survived into our times also pose puzzles. The centre of the Milky Way, and every galaxy like it, seems to host a black hole with a mass millions or even billions of times that of the sun. How did they get that big? One theory is that they began as star-sized black holes, produced when massive stars exploded and fell in on themselves, and then slowly grew by sucking in gas and surrounding stars. Yet a typical supermassive black hole would need longer than the age

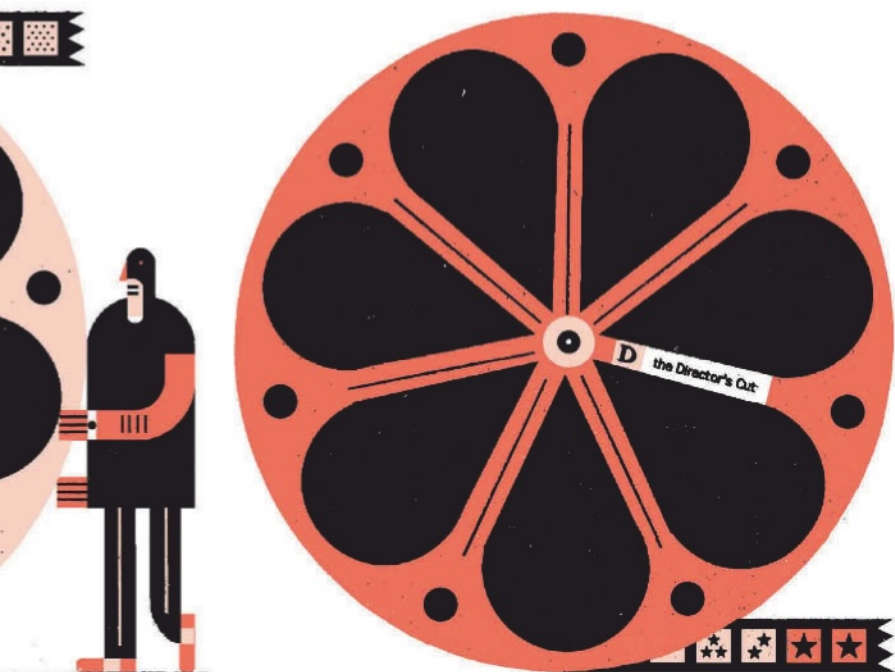
“What happened during the universe’s adolescent years has long been a matter of conjecture”

of the universe to swallow enough material. An alternative theory is that they were simply born big, produced directly by the collapse of massive amounts of primordial gas.

And was it ultraviolet light from the first stars, or X-rays emitted by these black holes as they fed, that caused the epoch of reionisation? Our Milky Way was a product of the universe’s dark ages – so asking these questions is again an investigation into our own origins.

We might look for answers by building even larger, more sensitive telescopes to look even further back towards the big bang. At least in the foreseeable future, however, that can only give us a partial view: the objects of interest are so far away that even the most gargantuan telescopes currently planned will see only the very brightest.

An alternative is to capture radio emissions from hydrogen atoms. Neutral hydrogen gas was an abundant, if diffusely spread, feature of the cosmos between the epochs of recombination and reionisation, and it gives off a faint signal of its presence. The lone electron and proton within each hydrogen atom act like two bar magnets that can lose energy by “flipping” so that their magnetic ➤



"Radio astronomy on Earth's surface is a bit like optical astronomy from the bottom of a swimming pool"

travels through it. The further away a region is, and so the earlier in cosmic time we see it, the more stretching the radiation undergoes. This allows us to map the ancient hydrogen in three dimensions: two across the sky and, by tuning our receivers to different stretched wavelengths, a third that is equivalent to distance or "look-back time".

The result is a film reel of the universe's missing years that can confirm – or, indeed, disprove – our general theoretical picture and answer many of those nagging questions. Fine details in the patterns of the 21-centimetre signal over time should reveal whether first-generation stars were a long-lasting feature in galaxy evolution, or brief candles that flickered once, never to be seen again. Different patterns of ionisation are expected from ultraviolet and X-ray emission, so that should tell us whether stars or black holes were the main agents of reionisation. If black holes were significant participants, the size of the ionised bubbles around them should reveal whether they were born big or ate their way to supersize.

Clues to a host of problems in fundamental particle physics might also be contained in the new hydrogen movies (see "Visible in radio", right). So why have we not made them before? Put simply, it has not been technically feasible. Hydrogen emissions from the first billion

moments, or spins, point in opposite directions. When an atom flips, it releases energy as a photon with a defined radio wavelength of 21 centimetres. Equally, a hydrogen atom can flip into the higher-energy, aligned state by absorbing a passing photon of the same wavelength.

Either way, the emission or absorption of 21-centimetre radiation over patches of sky is a sure sign that hydrogen atoms are present. Because hydrogen is ionised by high-energy radiation, whether from brightly burning stars or from supermassive galactic black holes, maps of where it is should provide a detailed picture of where stars and galaxies are not.

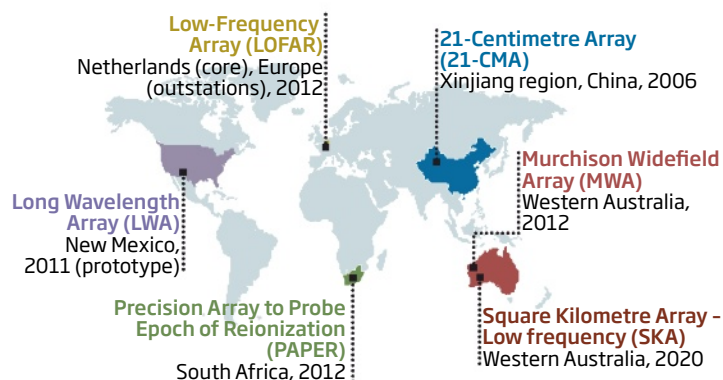
The hydrogen signal was predicted by Dutch astronomer Hendrik van de Hulst in 1942, and first picked up from the Milky Way in 1951 by Harold Ewen and Edward Purcell, who placed a horn antenna through a window of the physics department at Harvard University. Since then it has been used to detect warm hydrogen gas in our galaxy and others nearby – for example, to measure Doppler shifts to higher or lower wavelengths and so find out what parts of galaxies are moving towards and away from us.

A similar Doppler-like shift can be used to record the evolution of hydrogen in the early universe. As the universe expands over time, so does the wavelength of radiation that

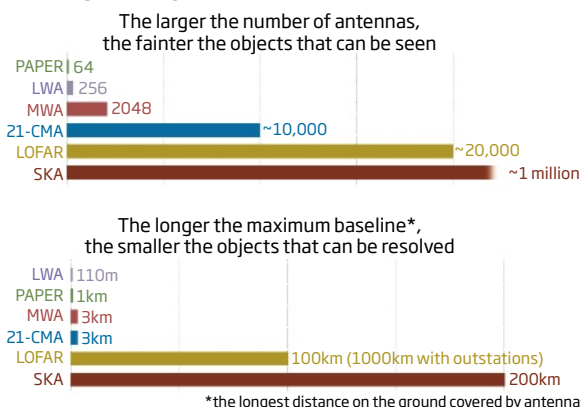
Radio multi-eyes

New arrays of radio antennas tuned to extremely long radio wavelengths will soon extend our vision back into the universe's crucial first billion years

Where are they?



How big are they?



years after the big bang are stretched to wavelengths of around 2 metres, several million times longer than a typical visible wavelength. The longer the wavelength, the larger the telescope required to capture it with the necessary resolution. For the sort of radio telescope that springs to most people's mind – an overgrown satellite dish – the size rapidly becomes unfeasibly large. The world's largest single radio telescope, a dish 305 metres in diameter built into a mountainside near Arecibo, Puerto Rico, does not come even close to the sensitivity required.

That is why the newest generation of radio telescopes takes a different approach. In an optical telescope, photons are generally separated by much more than their short wavelengths. The telescope is really just a bucket in which we collect individual photons to count them. In a radio telescope, by contrast, incoming photons from a distant source overlap and are recorded as a single continuous wave. This wave can be sampled at different points by many widely distributed small radio antennae rather than an analogue sound signal is sampled in time to make a digital recording, and the samples combined by computer algorithms into a single, coherent signal.

One such array, India's 30-antenna Giant Metrewave Radio Telescope, has been up and running since 1996, but has proved too small: it was designed in a data-starved era when we believed there was much more hydrogen in the universe than in fact there is. The more antennas within a given area, the greater the telescope sensitivity. The bigger the "baseline" of the telescope – the maximum distance on the ground covered by antennas – the smaller

VISIBLE IN RADIO

Hydrogen's 21-centimetre radio emissions should illuminate the universe's dark adolescence (see main story), but also address some deeper questions.

What was inflation?

The 21-centimetre observations ultimately measure variations in cosmic density seeded at the period of inflation, a breakneck expansion of space thought to have occurred a split second after the big bang. The cosmic microwave background radiation provides a single 2D projection of these fluctuations 400,000 years after the big bang. Observations at 21 centimetres will give us a far richer, 3D source of detail on what physics looked like under

extremely hot, dense conditions far beyond those recreated in earthly accelerators.

How massive are neutrinos?

That neutrinos have mass comes as a surprise to the standard model of particle physics – but they do, albeit a tiny and as yet unmeasured one. Characterising these elusive particles is important for understanding physics beyond the standard model, but it is a frustrating process using detectors on Earth. Measuring neutrinos' influence on how structures form in the universe is a surprisingly effective alternative. The passage of neutrinos smooths the distribution of matter, and the

scales on which this occurs at different times tells us how far and fast neutrinos have travelled since the big bang – and hence their mass.

What is dark matter?

Invisible dark matter is thought to make up 80 per cent of cosmic matter, and is needed to explain why galaxies rotate as fast as they do. Most particles postulated to be dark matter can annihilate each other, releasing energy and heating their surrounds. This would make an imprint in the 21-centimetre radiation, so measuring its pattern across the early cosmos will narrow down where dark matter was – and possibly reveal its nature.

the distant objects you can resolve. Imagine several thousand TV antennas connected to a supercomputer and you have an accurate image of a series of huge, long-wavelength radio arrays that are about to sharpen our view of the deep cosmic past (see diagram, below).

These arrays still won't have it easy. Earth's ionosphere interferes with radio waves as they pass, distorting the position and shape of distant sources and creating an effect that is like trying to perform optical astronomy from the bottom of a swimming pool. This must be corrected by referencing coordinates to a network of radio "pulsars", fast, rotating neutron stars that send regular, pinpointed

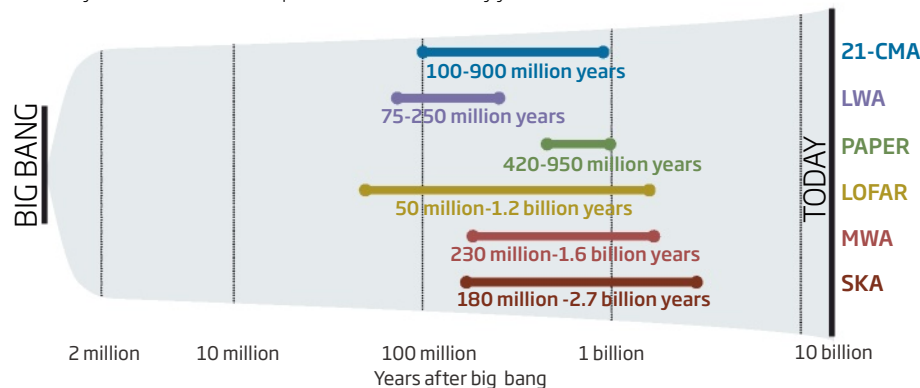
pulses of radiation our way. Individual bright radio sources and diffuse radio emissions from our own galaxy, which are 10,000 times brighter than the ancient cosmic signal, must also be peeled away – not to mention the busy noise of our own radio transmissions.

The theoretical and computational tools to overcome such difficulties are now largely in place, and the first definitive detection of hydrogen from the epoch of reionisation is expected within the next few years. These first pictures will still be a little fuzzy. For a finer-grained view, all eyes will be on the Square Kilometre Array (SKA), on which construction is due to start in South Africa and western Australia in 2018. The long-wavelength (low-frequency) part of the SKA, to be completed in Australia by about 2020, will consist of around one million radio antennas with a collecting area of a square kilometre, plugged into one of the world's fastest supercomputers. It should help to unravel the nature of the very first galaxies that formed in the first few hundred million years after the big bang.

Looking even further ahead, NASA is examining the possibility of 21-centimetre experiments on the far side of the moon, avoiding the problems both of Earth's ionosphere and of human radio activity. The idea is not as fanciful as it sounds. After all, in comparison to hulking a weighty conventional telescope with mighty mirrors into orbit, a simple radio array would be little more than a few wires connected to a supercomputer and a power source. Not much for the ultimate cinematic record of the universe's history. ■

How far back can they look?

Each array is tuned to a different span of the universe's early years



CHAPTER TWO
THE NATURE OF REALITY



Is material reality made of particles or waves? Take your pick, says Anil Ananthaswamy - but you'll be wrong

Quantum shadows



IF YOU haven't found something strange during the day," John Archibald Wheeler is said to have remarked, "It hasn't been much of a day." But then, strangeness was Wheeler's stock-in-trade. As one of the 20th century's leading theoretical physicists, the things he dealt with every day – the space- and time-bending warpings of Einstein's relativity, the fuzzy uncertainties and improbabilities of quantum physics – were the sort to boggle the minds of most mere mortals.

Even so, one day in 1978 must have been quite something for Wheeler. That was when he first hit on a very strange idea to test how photons might be expected to behave. Half a century earlier, quantum physics had produced the startling insight that light – everything in the quantum world, in fact – has a dual character. Sometimes it acts as if made of discrete chunks of stuff that follows well-defined paths – particles. At other times, it adopts the more amorphous, space-filling guise of a wave. That led to a question that exercised Wheeler: what makes it show which side, and when?

It took a while for the test Wheeler devised to become experimental reality. When it finally did, the answer that came was strange enough. Recently, though, the experiment has been redone with a further quantum twist. And it's probably time to abandon any pretence of understanding the outcome. Forget waves, forget particles, forget anything that's one or the other. Reality is far more inscrutable than that.

For centuries, light has illuminated our ideas of the material world. The debate about its nature, wave or particle, goes back to the philosophers of ancient Greece, and has featured luminaries such as Newton, Descartes and Einstein on one side or the other. By the dawn of the 20th century, the result was best described as a scoring draw, with both sides having gathered significant support (see "Duelling over duality", page 26).

The central mystery

Quantum physics broke the deadlock essentially by saying that everyone was right. The apparent proof comes with a quantum version of an experiment first performed by the English physicist Thomas Young in 1803, ironically to support the wave theory of light. Young shone light on a screen with two tiny, parallel slits in it. On another screen a distance behind the first, he saw alternating vertical fringes of light and dark that seemed incontrovertible proof of light's wave

character. Water waves passing through two narrow openings in a sea wall diffract and interfere in a similar way, sometimes constructively amplifying and sometimes destructively reducing each other beyond.

The strangeness starts when you lower the light intensity to the point at which only a single photon enters the experimental set-up at any one time. In 1905, Einstein had strongly suggested that a single photon is a particle, and indeed, place a detector at one or other of the slits and you hear the beep, beep of single particles hitting it. But remove the particle detector and place a light-collecting screen – a kind of long-exposure camera – a distance behind the slits, and the same pattern of light and shade that Young had observed slowly builds up. It is as if each photon is an

"The idea that physical reality depends on an observer's whim bothered the likes of Einstein no end"

interfering wave that passes simultaneously through both slits. The same happens with other quantum particles: electrons, neutrons, atoms and even 60-carbon-atom buckyballs.

For Niels Bohr, the great Danish pioneer of quantum physics, this "central mystery" was nothing less than a principle of the new theory, one he called the complementarity principle. Quantum objects such as photons simply have complementary properties – being a wave, being a particle – that can be observed singly, but never together. And what determines which guise an object adopts? Bohr laid out a first outline of an answer at a grand gathering of physicists at the Istituto Carducci on the shores of Lake Como in Italy in September 1927: we do. Look for a particle and you'll see a particle. Look for a wave and that's what you'll see.

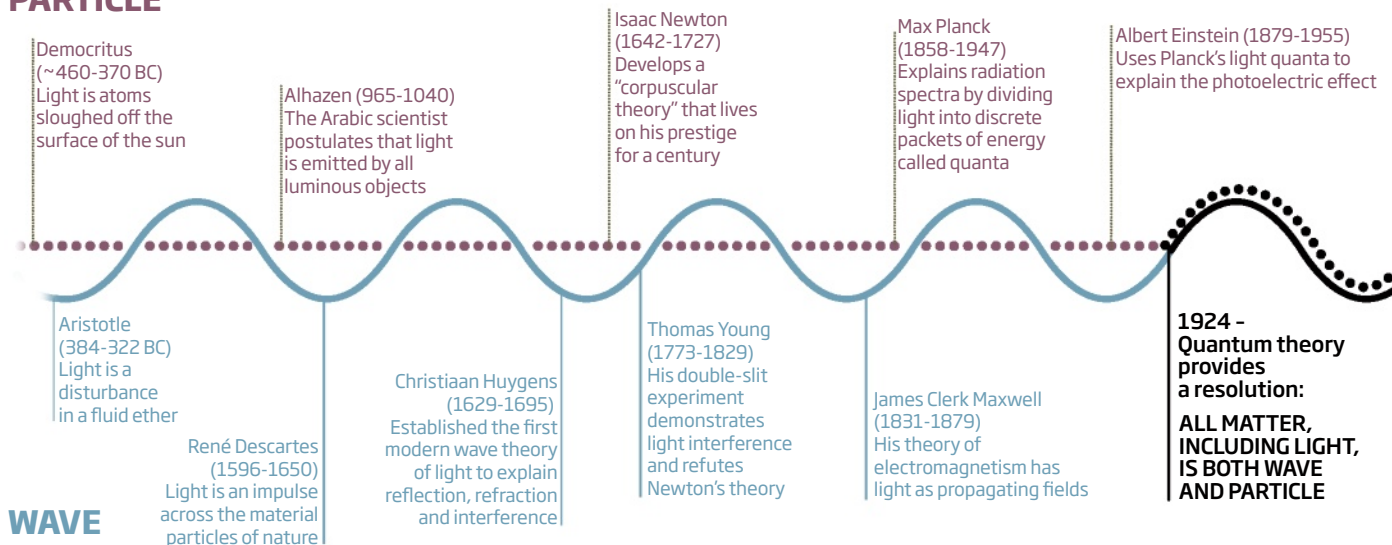
The idea that physical reality depends on an observer's whim bothered the likes of Einstein no end. "No reasonable definition of reality could be expected to permit this," he huffed in a famous paper he co-authored in 1935 with Boris Podolsky and Nathan Rosen. Einstein favoured an alternative idea of an underlying but as yet inaccessible layer of reality containing hidden influences that "told" the photon about the nature of the experiment to be performed on it, changing its behaviour accordingly.

There is more to this than wild conspiracy theory. Imagine an explosion that sends two ➤

Duelling over duality

Philosophers and physicists have flip-flopped on whether light is a wave or a particle all the way back to ancient Greece

PARTICLE



pieces of shrapnel in opposite directions. The explosion obeys the law of conservation of momentum, and so the mass and velocity of the pieces are correlated. But if you know nothing of momentum conservation, you could easily think that measuring the properties of one fragment determines the properties of the other, rather than both being set at the point of explosion. Was a similar hidden reality responsible for goings-on in the quantum world?

This is where Wheeler's thought experiment came in. Its aim was to settle the issue of what told the photon how to behave, using an updated version of the double-slit experiment. Photons would be given a choice of two paths to travel in a device known as an interferometer. At the far end of the interferometer, the two paths would either be recombined or not. If the photons were measured without this recombination – an "open" interferometer – that was the equivalent of putting a detector at one or other of the slits. You would expect to see single particles travelling down one path or the other, all things being equal, splitting 50:50 between the two (see "Neither one nor the other", opposite).

Alternatively, the photons could be measured after recombination – a "closed" setting. In this case, what you expect to see

"Wave and particle are two sides of one coin describing reality. As to which way it flips - you decide"



depends on the lengths of the two paths through the interferometer. If both are exactly the same length, the peaks of the waves arrive at the same time at one of the detectors and interfere constructively there: 100 per cent of the hits appear on that detector and none on the other. By altering one path length, however, you can bring the wave fronts out of sync and vary the interference at the first detector from completely constructive to totally destructive, so that it receives no hits. This is equivalent to scanning across from a bright fringe to a dark one on the interference

screen of the double-slit experiment.

Wheeler's twist to the experiment was to delay choosing how to measure the photon – whether in an open or a closed setting – until after it had entered the interferometer. That way, the photon couldn't possibly "know" whether to take one or both paths, and so if it was supposed to act as a particle or a wave.

Or could it?

It was almost three decades before the experiment could be done. To make sure there was no hidden influence of the kind favoured by Einstein, you needed a very large interferometer, so that no word of the choice of measurement could reach the photon, even if the information travelled at light speed (anything faster was expressly forbidden by Einstein's own theory of relativity). In 2007, Alain Aspect and his team at the Institute of Optics in Palaiseau, France, built an interferometer with arms 48 metres long. The result? Whenever they chose at the last instant to measure the photons with a closed interferometer, they saw wave interference. Whenever they chose an open interferometer, they saw particles.

There was no getting round it. Wave and particle behaviours really do seem to be two sides of one coin representing material reality. As to which way it flips – well, you decide. "Isn't that beautiful?" said Aspect in a public

lecture at the Physics@FOM conference in Veldhoven, the Netherlands, in 2012. "I think there is no other conclusion to draw from this experiment."

Unless, of course, you make things even stranger. In December 2011, Radu Ionicioiu of the National Institute of Physics and Nuclear Engineering in Bucharest, and Daniel Terno of Macquarie University in Sydney, Australia, proposed extending Wheeler's thought experiment. Their new twist was that the decision of how to measure the photon, as a particle or as a wave, should itself be a quantum-mechanical one – not a definite yes or no, but an indeterminate, fuzzy yes-and-no.

Infinite shades of grey

There is a way to do that: you use light to control the detector designed to probe the light. First you prepare a "control" photon in a quantum superposition of two states. One of these states switches the interferometer to an open, particle-measuring state, and the other to a closed, wave-measuring state. Crucially, you only measure the state of the control photon after you have measured the experimental "system" photon passing through the interferometer. As far as you are concerned, the system photon is passing through an interferometer that is both open and closed; you don't know whether you are setting out to measure wave or particle behaviour. So what do you measure?

This time, it took only a few months for the experimentalists to catch up with the theorists. But when three independent groups, led by Chuan-Feng Li at the University of Science and Technology of China in Hefei, Jeremy O'Brien at the University of Bristol, UK, and Sébastien Tanzilli at the University of Nice, France, performed different versions of the experiment in 2012, the results were unnerving – even to those who consider themselves inured to the weirdnesses of quantum physics.

The answer is, what you see depends on the control photon. If you look at the measurements of the system photons without ever checking the corresponding measurements of the control photons – so never knowing what measurement you made – you see a distribution of hits on the two detectors that is the signature neither of particles or waves, but some ambiguous mixture of the two. If particle is black and wave is white, this is some shade of grey.

Do the same, but this time looking at the control-photon measurements as well, and it

is like putting on a pair of magic specs. Grey separates clearly into black and white. You can pick out the system photons that passed through an open interferometer, and they are clearly particles. Those that passed through a closed interferometer look like waves. The photons reveal their colours in accordance with the kind of measurement the control photon said you made.

It gets yet stranger. Quantum mechanics allows you to put the control photon not just in an equal mix of two states, but in varying proportions. That is equivalent to an interferometer setting that is, say, open 70 per cent of the time and closed 30 per cent of the time. If we measure a bunch of system photons in this configuration, and look at the data before putting on our magic specs, we see an ambiguous signature once again – but this time, its shade of grey has shifted closer to particle black than wave white. Put on the specs, though, and we see system photons 70 per cent of which have seemingly, but clearly, behaved as particles, while the remaining 30 per cent acted as waves.

In one sense, the results leave Bohr's side of

the argument about quantum reality stronger. There is a tight correlation between the state of the control photon, representing the nature of the measurement, and the system photon, representing the state of reality. Make for more of a particle measurement, and you'll measure something more like a particle, and vice versa. As in earlier experiments, a hidden-reality theory à la Einstein cannot explain the results.

But in another sense, we are left grappling for words. "Our experiment defies the conventional boundaries set by the complementarity principle," says Li. Ionicioiu agrees. "Complementarity shows only the two ends, black and white, of a spectrum between particle and wave," he says. "This experiment allows us to see the shades of grey in between."

So, has Bohr been proved wrong too? Johannes Kofler of the Max Planck Institute of Quantum Optics in Garching, Germany, doesn't think so. "I'm really very, very sure that he would be perfectly fine with all these experiments," he says. The complementarity principle is at the heart of the "Copenhagen interpretation" of quantum mechanics, named after Bohr's home city, which essentially argues that we see a conflict in such results only because our minds, attuned as they are to a macroscopic, classically functioning cosmos, are not equipped to deal with the quantum world. "The Copenhagen interpretation, from the very beginning, didn't demand any 'realistic' world view of the quantum system," says Kofler.

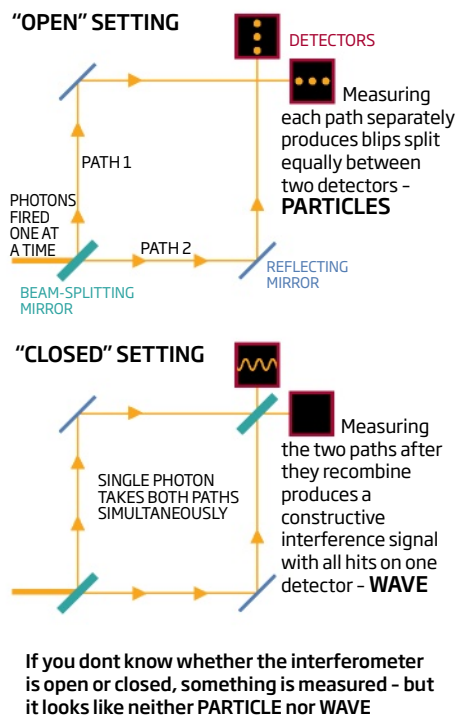
The outcomes of the latest experiments simply bear that out. "Particle" and "wave" are concepts we latch on to because they seem to correspond to guises of matter in our familiar, classical world. But attempting to describe true quantum reality with these or any other black-or-white concepts is an enterprise doomed to failure.

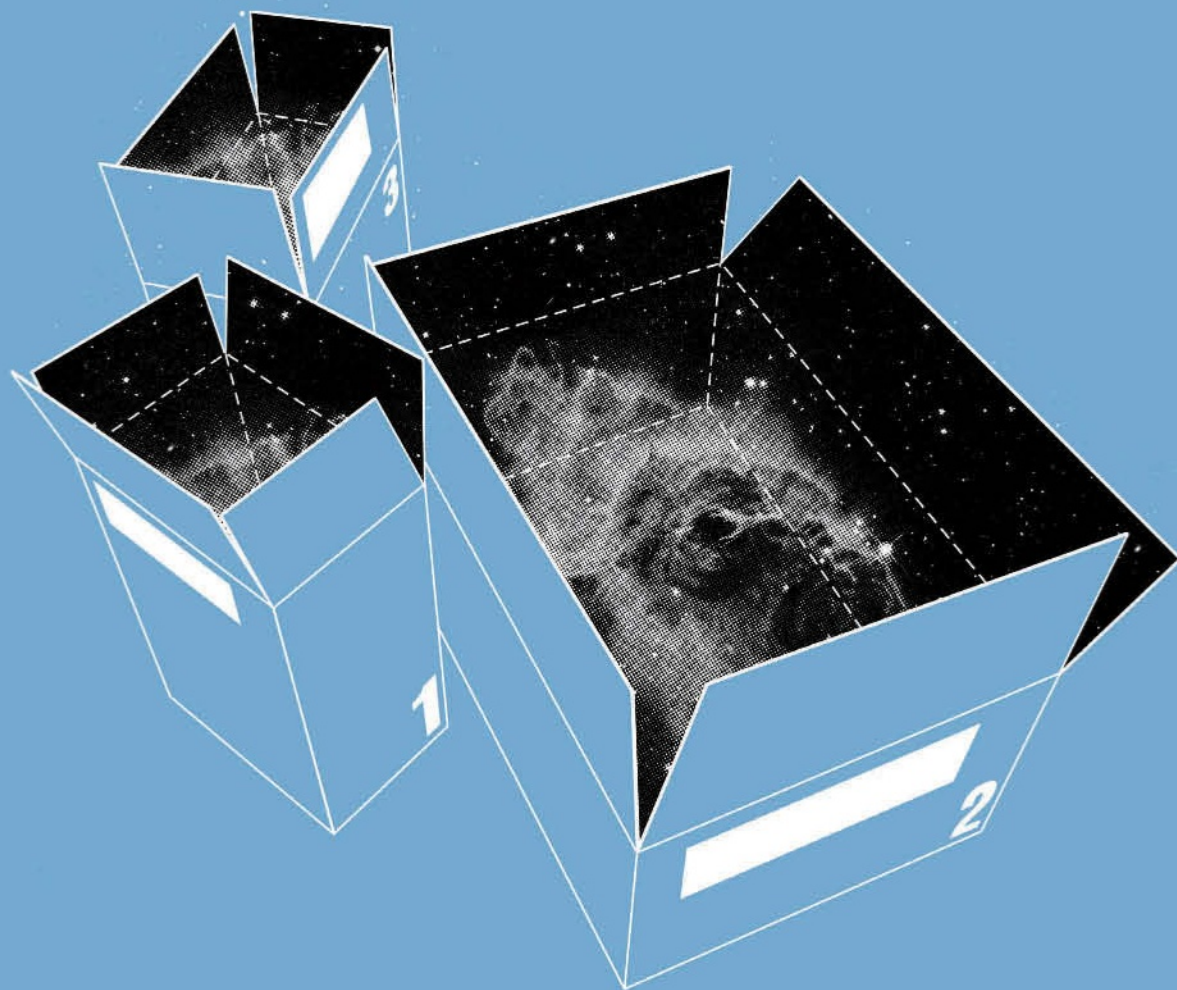
It's a notion that takes us straight back into Plato's cave, says Ionicioiu. In the ancient Greek philosopher's allegory, prisoners shackled in a cave see only shadows of objects cast onto a cave wall, never the object itself. A cylinder, for example, might be seen as a rectangle or a circle, or anything in between. Something similar is happening with the basic building blocks of reality. "Sometimes the photon looks like a wave, sometimes like a particle, or like anything in between," says Ionicioiu. In reality, though, it is none of these things. What it is, though, we do not have the words or the concepts to express.

Now that is strange. And for quantum physicists, all in a day's work. ■

Neither one nor the other

According to how it is set up, an interferometer can be used to "prove" light is particles, waves – or nothing of the sort





SEEING TRIPLE

Space has three dimensions. Why, asks *Matthew Chalmers*

THE ONLY way is up." An earnest student of our physical realities might find room to dispute this jolly phrase. There is also down, and, for that matter, left, right, forwards and backwards. Six ways to go. Then again, the further up you go, the less down you are, and similarly for left and right, forwards and backwards. So that's three independent directions to move in – gravity and local obstacles permitting.

It is a fact so bald that we rarely stop to ask an even balder question: why?

Physicists have wrestled with this perplexing question of space's essential three-ness for a good while now – not, it must be said, with much success. Our best theories of nature supply no clue as to why space might have three dimensions, rather than two, four or 5.2. Even worse, the drive for ever-grander replacements keeps finding hints that the magic number is anything but three.

Perhaps we have been trying too hard. The latest attempt to explain the mystery of three suggests an answer is under our nose, hidden in the workings of physics we already know.

For a long while, our understanding of the world progressed rather nicely by ignoring space. Newton first formalised the concept in 1687 when he introduced his theory of gravity. For him, space and time were real enough, but no more than an impassive backdrop against which more interesting stuff such as falling

apples and orbiting planets happened. It was Einstein who forced us to buck up our ideas. In his general theory of relativity, formulated in 1916, space becomes a dynamic entity. Interwoven with time into a 4D space-time, it is bent and buckled by matter to create the force we call gravity.

Oddly, though, general relativity says nothing about space's most obvious quality, its three-ness. With a little tweaking, relativity's mathematics works fine in any number of dimensions.

Some have filled the gap with "anthropic" reasoning: space must be just so for us to pose questions about it (see "We're here because we're here", page 31). Others think the answer is there within physics; we just haven't found it yet. A theory of quantum gravity would unify general relativity with quantum mechanics. It would be valid in the first, searing moments of the universe when space-time was itself a tiny, jumbled quantum mess. Perhaps the three-dimensionality of space, not to mention the one-dimensionality

of time, emerged from physics unknown during this hectic era. "Emergent space-time is the modern view," says Gary Gibbons of the University of Cambridge.

It is a difficult one to test. Vast machines such as the Large Hadron Collider at CERN near Geneva, Switzerland, smash particles together at humongous energies to push back our understanding of physics towards those energetic first few moments, but they are still a way off. So far, no experiment has produced any lead in the problem of quantum gravity.

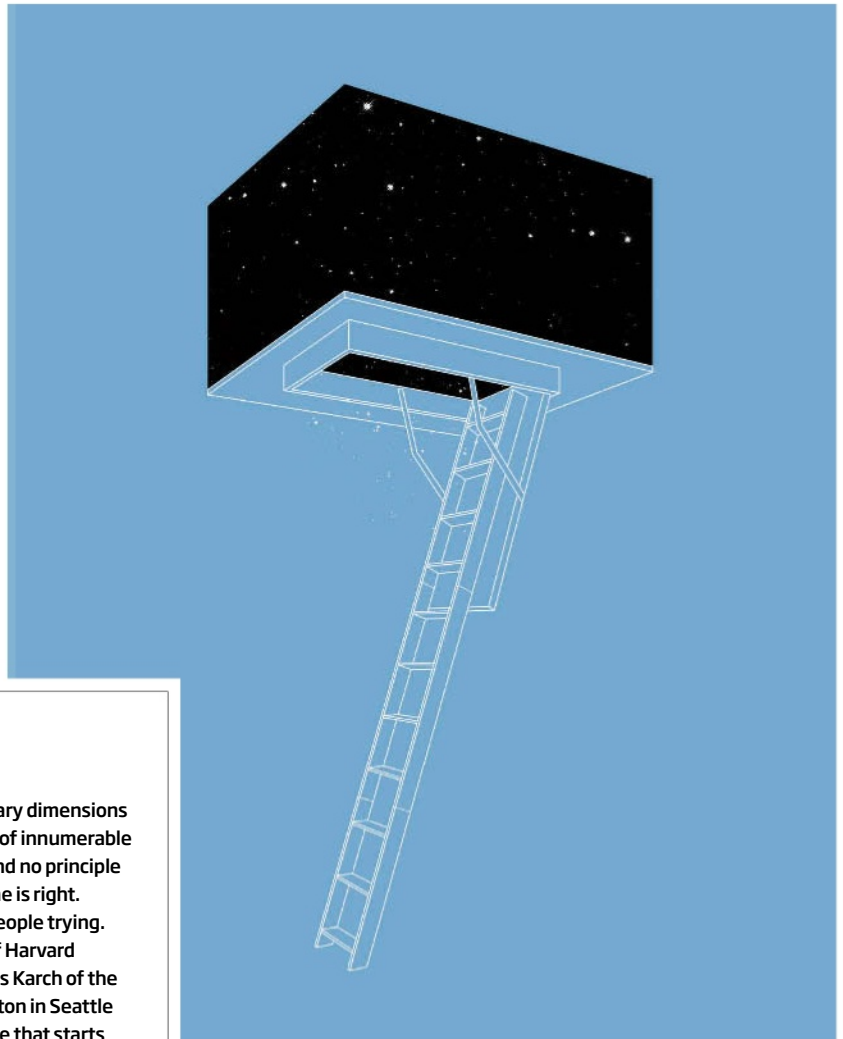
As far as the mystery of three is concerned, it doesn't help that the most popular theoretical ideas seem to work in anything other than 3D. String theory, for instance, requires at least six extra spatial dimensions (see "This one goes up to eleven", page 30). Causal dynamical triangulations, an alternative approach, depicts a quantum network of 2D elements that evolves into a macroscopic 3D space. Loop quantum gravity, meanwhile, replaces a smooth space-time with a version that is grainy or foamy at the smallest scales – but it takes the number 3 as an input and has done with it, says one of the theory's originators, Carlo Rovelli of the University of Aix-Marseille in France. "This just does not seem to be at all one of those questions that science has tools for figuring out at the moment," he says.

Markus Müller begs to differ. A theoretical ►

"It doesn't help that most grand theories of nature seem to work in anything other than 3D"

physicist at the University of Heidelberg in Germany, he thinks we can solve the question of three – and he reckons the answer lies buried in the roots of a theory we already have. Quantum theory certainly describes the physical world phenomenally well, but it is offensive to our conventional views of reality, allowing objects to be in two states or places at once, for instance, and generally denying principles we hold dear, such as clear lines of cause and effect. The traditional response of physicists has been to ignore these little local difficulties – known as the “shut up and calculate” approach – or to tie themselves in knots with all manner of interpretations, from strange influences of experimenters on experiments to the existence of many bifurcating worlds.

Müller had an idea to sidestep these “philosophical pitfalls”, as he calls them: derive a basis for quantum mechanics from examining what experience allows. More



THIS ONE GOES UP TO ELEVEN

String theory says the fundamental units of reality are not point-like, zero-dimensional particles, but one-dimensional strings. The tricky bit is that these strings need at least nine spatial dimensions to move in to preserve the theory's mathematical consistency. The extra six are assumed to be “compactified” at scales of a billionth of a trillionth of a trillionth of a centimetre.

At least, that was the old-fashioned view. These days, the very notion of dimensionality in string theory is vague, says theorist Cumrun Vafa of Harvard University. In an overarching 11D mega-idea called M-theory, physics plays out on mathematical objects called branes that have an arbitrary number of dimensions fewer than 10.

In this picture, our universe could be bound to a “3-brane” within a higher-dimensional space, floating around with numerous others that are invisible, but right in front of our eyes. Or perhaps not; perhaps our 3D theories are holograms hailing from some higher-dimensional structure. The many different ways to play with

string theory's imaginary dimensions present a “landscape” of innumerable potential universes, and no principle to determine which one is right.

That doesn't stop people trying. In 2005 Lisa Randall of Harvard University and Andreas Karch of the University of Washington in Seattle showed how a universe that starts out with equal numbers of branes and anti-branes – branes with opposite orientations – would evolve into one dominated by 3-branes and 7-branes. Vafa's own work with Robert Brandenberger of McGill University in Montreal, Canada, suggests that the dimensionality of a string-theory universe is intimately entwined with its expansion: if the universe began as an unimaginably small 9D ball of string that suddenly began to unwind, three would become infinitely large and the rest would remain curled up.

That idea was revived in 2012 by Brian Greene of Columbia University in New York and his colleagues. Even so, Vafa admits the case is not closed. “Right now, we cannot claim that string theory predicts the dimensionality of space,” he says.

than a century ago, we similarly gained insight into the laws of thermodynamics by observing what could and couldn't happen in nature – you can't have a perpetual-motion machine that gives you energy for nothing, for example. “One thinks very carefully about what the formulas mean, and what can actually be done concretely or operationally,” says Müller.

Working with Lluís Masanes of the University of Bristol in the UK, Müller looked at a situation where a sender and receiver set out to exchange information encoded in quantum states – the basis of the very real, super-secure technique of quantum cryptography. Countless experiments have shown that quantum information held by the sender and receiver is correlated above and beyond what is possible in the classical world. Change the quantum state of photons used to

encode the message at one end, say, and you might see spooky, instantaneous changes in the states of photons at the other end.

The duo started with a few “reasonable” assumptions about how the physical universe surrounding such a sender and receiver worked. It must have a certain number of dimensions, including one of time, for example, and there must be some way for information to flow within it. In addition, they assumed that at least some of the physical processes in the world worked randomly, although they did not specify how randomly.

What follows is a dense mathematical proof with a surprising outcome. Not only is quantum theory the only theory that can supply the degree of randomness and correlation seen in nature, but it can only do so if space is 3D, too.

Roots of reality

That could be just a mathematical coincidence. Quantum states are described not by 1D real numbers, which all lie on a single line, but by 2D complex numbers that represent points on a plane. The way these numbers interact to produce a complete description of objects that can be in more than one state at once, such as photons, naturally sketches out a 3D sphere describing all those possible states. Perhaps this result is just emphasising how the dimensionality of basic quantum objects and that of space happen to be the same.

Müller thinks not: he thinks it points to an inextricable link between space’s geometry and the degree of probability inherent in quantum theory. If so, the roots of relativity and quantum theory would be embedded in the way information is exchanged in the cosmos, suggesting where to look for any unifying theories. “It offers a clue that the notion of information will be an important part of quantum gravity,” says Müller.

Since his and Masanes’s idea was published, Borivoje Dakic and Caslav Brukner at the University of Vienna, Austria, have used similar arguments from information theory to restrict the dimensionality of space. They, too, show that only in a 3D universe can quantum mechanics hold true – at least in a universe where microscopic objects interact “pairwise” with each other, as they appear to in ours. Relax this restriction so that three or more quantum systems can interact at the same time, and higher-dimensional universes are just fine.

Dorje Brody of Brunel University in London thinks we must treat such arguments with

WE’RE HERE BECAUSE WE’RE HERE

Life in fewer than three dimensions would be unimaginably different. Our minds are made of neurons criss-crossing in 3D, cylindrical tubes provide our digestive tracts and 3D helices transmit our genetic blueprints. Two-dimensional life might be too limited to even ask questions about the dimensionality of space.

Other arguments might discount higher numbers. In our world of three spatial dimensions, the gravitational force between two objects falls off with the square of the distance between them. In two dimensions, it would decrease far slower, in proportion to the distance. In a 4D world, it would fall off more steeply as the cube of the distance.

In both cases, planetary orbits would be unstable: the merest asteroid strike would be enough to cause a planet to spiral into its parent star or fly out into deep space. Not that there would be planets or stars – similar stability arguments apply to the electromagnetic force that holds electrons in orbit around atomic nuclei.

The dimensionality of space, then, is one of those numbers in physics, like the constants that determine the strength of electromagnetism and gravity, that dictate whether complex life can evolve. “In my opinion it is the only known explanation,” says physicist Leonard Susskind of Stanford University in California. Atoms and molecules would not exist in the form that we know them in anything other than three dimensions.

Others are unimpressed by such arguments. “They are nonsense,” says Carlo Rovelli of the University of Aix-Marseille in France. “If we lived in six dimensions, many people would find equally good arguments to show that only six dimensions are possible.” Critics also point out that, unlike constants of nature, the dimensionality of space is a small, perfect integer. It is hard to imagine something that neat and rounded dropping out of some random environmental process.

care. They start out with a few assumptions that seem indisputable, but additional ones are often “smuggled in” further on, he says. “Somewhere along the line the rabbit has to enter the hat.”

He thinks the conclusions we draw depend critically on the mathematical language we use. There are consistent representations of quantum mechanics using not 2D complex numbers, but 4D “quaternions” and 8D “octonions”. Brody and his colleague Eva-Maria Graefe of Imperial College London have shown that in the quaternion formulation, five dimensions of space naturally drop out, whereas octonion quantum mechanics demands nine.

Back to square one? Not quite. These reformulations predict subtly different outcomes for certain experiments than does conventional quantum mechanics, including different degrees of correlations between particles. If quantum interactions truly are the source of space’s dimensionality, that suggests new ways to pin the answer down in the lab without the huge expenditure on machines such as the Large Hadron Collider, says Brody.

Terry Rudolph, also of Imperial College, thinks it is an intriguing avenue to investigate, but cautions against jumping on the latest bandwagon. “Remember when chaos theory

“The roots of relativity and quantum theory might lie in the way information is exchanged in the cosmos”

was meant to solve anything and everything?” he says. His hunch is that the dimensionality of space might be too human a concept for us to explain – a view articulated by the philosopher Immanuel Kant in the 18th century. Kant described space as “a subjective condition of sensibility” whose only significance was in allowing us to specify the relationship between other objects. “Perhaps it is just a variable that these particular monkeys evolved to find useful in their quest for bananas,” says Rudolph.

Newton and Einstein, both proponents of the reality of space, presumably would have approved of the ongoing attempts to explain something so basic as its three-ness. In the end, though, we might be forced to accept it as a brute fact, says philosopher Craig Callender of the University of California in San Diego. “It might be an unexplained explainer.” ■



It's looking likely our universe is just one of many. But what kind of multiverse do we live in, asks **Robert Adler**

The many faces of the multiverse

WHETHER we are searching the cosmos or probing the subatomic realm, our most successful theories lead to the inescapable conclusion that our universe is just a speck in a vast sea of universes.

Until recently many physicists were reluctant to accept the idea of this so-called multiverse. Recent progress in cosmology, string theory and quantum mechanics, though, has brought about a change of heart. "The multiverse is not some kind of optional thing, like can you supersize or not," says Raphael Bousso, a theoretical physicist at the University of California, Berkeley. Our own cosmological history, he says, tells us that "it's there and we need to deal with it".

These days researchers like Bousso are treating multiverses as real, investigating them, testing them and seeing what they say about our universe. One of their main motivations is the need to explain why the physical laws underlying our universe seem so finely tuned as to allow galaxies, stars, planets, complex chemistry, life – and us – to exist. Rather than appealing to God or blind luck, some argue that our existence sets parameters that reliably pluck universes like ours from the bottomless grab-bag of the multiverse.

Yet there is a problem. Different theories spin off very different kinds of multiverses. Our current standard theory of how the universe came to be, for example, predicts

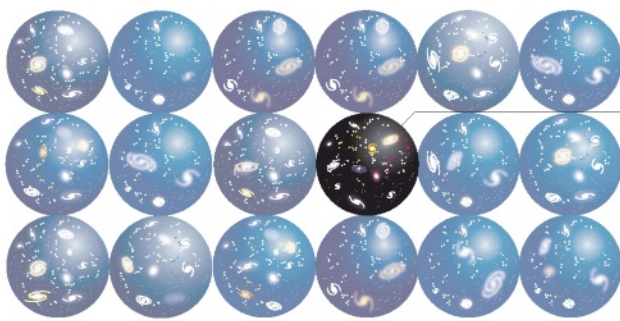
an infinite expanse of other universes, including an infinite number in which duplicates of you are reading this sentence and wondering if those other versions of you really exist. Meanwhile, string theory, which hoped to derive the particles, forces and constants of our universe from fundamental principles, instead discovered a wilderness of 10^{500} universes fundamentally different from ours. Even quantum mechanics implies that our universe is a single snowflake in a blizzard of parallel universes (see diagrams, "The multiverse hierarchy", on following pages).

We are now faced with trying to understand and relate these ideas to one another. Of late, we have made enormous strides in making theoretical sense of concepts of the multiverse. At the same time, several groups claim to have made astronomical observations in support of the idea. It seems that we are finally beginning to find our place within the universe of universes.

We can trace the idea of the multiverse back to the early 1980s, when physicists realised that the big bang created an equally big problem. When astronomers measured its afterglow, radiation known as the cosmic microwave background (CMB), they found that it was unfathomably uniform – even at opposite ends of the visible universe. Finding a temperature match between such widely separated regions to within 1/10,000th of a ➤

The multiverse hierarchy: level 1

If the big bang started with a period of inflationary growth, there would be a multitude of universes a lot like ours – but with different arrangements of matter



Our universe
42 billion light years across – the distance light has travelled in our expanding universe

Physics:

Like ours, but with all possible initial conditions and histories replicated an infinite number of times

Support:

Plays to the idea of the principle of mediocrity – that there's nothing special about the universe we see

Relationships:

All level 1 universes bear a family resemblance to ours and to each other

Connections:

Since everything that can happen in our universe has happened in some other level 1 universe, there may be a direct connection between level 1 and level 3 quantum multiverses

degree, as we now know it to be, was as surprising as finding that their inhabitants, if any, spoke the same language, says Brian Greene, a physicist at Columbia University in New York City.

The problem was solved brilliantly by cosmologists Alan Guth at the Massachusetts Institute of Technology and Andrei Linde at Stanford University, California, among others. Their insight was that in the first 10^{-35} seconds of the universe's existence, space itself expanded by a factor of roughly 10^{30} . This stupendous stretching, known as inflation, accounts for the uniform temperature of the CMB and resolves another conundrum: why space appears flat, like a three-dimensional version of an infinite tabletop. Inflation has become an incredibly successful theory, precisely predicting the subtle ripples since measured in the CMB,

which are echoes of quantum perturbations thought to have seeded galaxies and stars.

Guth and Linde's attempt to explain our universe also led directly to a multiverse. That's because inflation didn't conveniently stop at the farthest regions from which light can travel to us today. Depending on how inflation unfolded, Guth says, the universe could be 10^{10} times, 10^{20} times, or even infinitely larger than the region we see. Inflation implies an expansion faster than the speed of light, meaning that beyond the horizon of our observable universe lie other parts of the universe that are effectively separated from ours. No influence can travel between these regions, essentially creating an infinite number of other worlds.

What would they look like? Max Tegmark, also a cosmologist at MIT, points out that although inflation predicts an abundance of universes, they all feature the same particles, forces and physical laws we find in our cosmic patch. But while in our universe elementary particles come together to make stars and galaxies in a particular way, the universe next door will contain a different arrangement of stars and galaxies, and so will our neighbours' neighbours. Still, Tegmark has shown that even if a universe like ours were completely filled with elementary particles, they can only be arranged in a finite number of ways. It's a huge number, 2 to the power of 10^{118} , but since there's no sign that space is finite, there is room for every arrangement to repeat.

This means that if you travel far enough, you will eventually encounter a universe containing an identical copy of you. Tegmark

has calculated that your nearest copy lives about 10 to the power of 10^{29} metres away. Carry on and you will find our universe's twin lies 10 to the power of 10^{118} metres from here. Since an infinite universe hosts an infinite number of variations, somewhere you have just won the Olympic 100 metres gold medal. Congratulations!

Abundant as these universes are, there is nothing exotic about them. Tegmark classes the universes implied by simple inflation or an infinite expanse of space as the first level of a four-tier hierarchy that gets much, much stranger.

Take the second type of multiverse. Soon after inflation was discovered, Linde realised that it could be an ongoing or eternal process. When the enormous energy of empty space creates an inflating baby universe, the space around it, still crackling with energy, continues to expand even faster. That space could then sprout more universes that themselves inflate, and so on and on. "Practically all models of inflation that have been discussed predict eternal inflation," says Alexander Vilenkin at Tufts University in Medford, Massachusetts, who pioneered the idea in the 1980s. Guth dubs it the ultimate free lunch.

The eternal-inflation smorgasbord includes an infinite number of level 1 universes, but many other varieties as well. Each universe turns out in a different way, so features once thought universal, such as the masses of elementary particles and the strength of fundamental forces, differ. The bubble universes from eternal inflation include every permutation the laws of physics allow, leading Linde to quip that it's not just a free lunch but "the only one at which all possible dishes are available".

Those universes are part of the second level of Tegmark's multiverse hierarchy. It also includes some 10^{500} sorts of universe implied by string theory, the leading contender for a "theory of everything" that would explain all the particles and forces of nature.

Today's standard model of particle physics includes a score of parameters whose values physicists can measure but can't explain, such as the mass of the electron. String theorists hoped their theory would explain why those parameters have the values they do, and so why our universe is the way it is.

They were sorely disappointed. Rather than producing one perfect snowflake – the particles, forces and interactions underpinning our universe – string theory loosed an avalanche of universes, a daunting expanse that Leonard Susskind, a theoretical



physicist at Stanford University, dubbed the string theory landscape.

What sets these universes apart is the nature of their space-time. In string theory, nature's particles and forces come about from vibrations of tiny strings in 10 dimensions. The reason we only experience four dimensions is because the rest are "compactified" or knotted into intricate structures too small for us to experience. The physics that plays out in any given universe depends on how many dimensions are crunched up, and their structure. Researchers have identified an enormous number of possible shapes that interact with string-theory fields to define a vast number of universes, most with unfamiliar physical laws and radically different forces and particles.

Eternal inflation provided a convincing mechanism for populating every point in the string theory landscape with an infinite number of real universes. "Originally, string theorists did not like the idea of the multiverse and wanted to have just one solution, but instead they found 10^{500} ," says Linde. "Now we must learn how to live in a universe which offers us an enormous multitude of choices."

Finding out why our universe is as it is, when there is such a vast number of alternatives, remains one of cosmology's

"Far from explaining why our universe is the way it is, string theory loosed an avalanche of alternative universes"

biggest challenges. Our universe seems inexplicably finely tuned to produce the conditions necessary for life. If gravity were a bit stronger, the big bang would have been a squib; a bit weaker and it couldn't have made galaxies and stars. If the electron's charge differed by a few per cent, stars could not have created the heavy elements that make Earth-like planets. If the strong force varied by half a per cent, carbon wouldn't exist, so neither would life as we know it. Topping the fine-tuning list is our universe's small cosmological constant – the tiny dose of dark energy that is the source of the accelerating expansion of the universe.

The discovery in the late 1990s that the universe's expansion is accelerating shocked most cosmologists. Quantum theory predicts a level of dark energy roughly 10^{120} times larger than what was found. Since that would blast the universe apart, most researchers had

assumed that some undiscovered symmetry would cancel that huge number, leaving a cosmological constant of zero. Nobody predicted that it would not be zero – except one person.

A decade earlier, Steven Weinberg, a Nobel-prize-winning physicist at the University of Texas, Austin, had predicted a small positive cosmological constant. That coup came from applying anthropic reasoning to the multiverse, an approach that is still hotly contested. Weinberg reasoned that for a universe to generate galaxies – and so stars, planets and observers – the amount of dark energy had to fall within a certain range for us to be here to measure it. This amounts to homing in on a subset of universes within the multiverse that have those properties. This probabilistic approach allowed Weinberg to predict the value of the cosmological constant with remarkable accuracy.

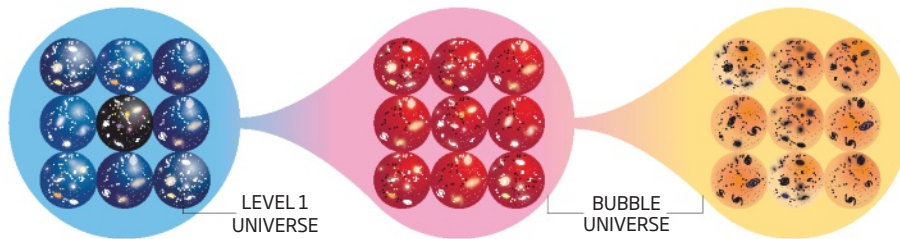
"The discovery of the cosmological constant was one of the most unexpected discoveries of the last century, and it was predicted by the multiverse," says Vilenkin. "So it's indirect evidence that we are living in a multiverse."

Since then, other researchers have used anthropic reasoning to constrain the amount of dark energy, the ratio of dark matter to ordinary matter, and the mass of elementary particles such as neutrinos and quarks. Using anthropic reasoning to winnow out our kind of universe from the multiverse, it seemed, might explain that mysterious fine-tuning. It might, in fact, be the only way. "If the laws and physical constants are different in other places," says Martin Rees, an astronomer at the University of Cambridge, "there's no avoiding it."

Unfortunately there's a catch in using this approach to elucidate our universe's place in the multiverse: the usual rules of probability may not apply, making it impossible to estimate the likelihood of universes like ours. "You have infinitely many places where you win the lottery, and infinitely many where you don't. So what's your basis for saying that winning the lottery is unlikely?" says Bousso. "It pulls the rug out from under us to prove a theory right or wrong."

The multiverse hierarchy: level 2

In the theory of eternal inflation, the space between universes continues to expand, and a limitless number of new "bubble" universes, with very different properties, continue to form. Each bubble universe contains an infinite number of its own level 1 universes



Physics:

Other bubble universes exhibit different laws of physics and have different dimensionality, particles, constants and forces to those seen in our universe. We might eventually discover that all these parameters flow from the same "theory of everything"

Support:

Inflation explains the uniformity and flatness of our universe and details of the cosmic microwave background. Eternal inflation implies bubble universes and provides a way of supplying string theory with the many universes it demands

Relationships:

Level 2 universes vary greatly. They represent separate bubbles or domains with different properties, and are separated from each other by inflating space

Connections:

Level 2 includes all possible level 1 universes plus an enormous variety of much stranger universes. Since everything that can happen in a particular level 2 universe has happened in other universes, level 2 may also correspond to the universes in the level 3 quantum multiverse

"One interpretation of quantum theory leads to the existence of a vast swarm of universes paralleling ours, in which all the theory's possibilities play out"



This "measure problem" may have been solved by Bousso and I-Sheng Yang of the University of Amsterdam in the Netherlands. They got rid of the troublesome infinities by deriving probabilities for a universe's parameters from local "causal patches" – everything an observer can ever interact with. The probabilities they found matched those from an alternative approach pioneered by Vilenkin and Jaume Garriga, at the University of Barcelona in Spain, putting predictions from the multiverse on what could be a solid footing for the first time.

While theorists have made great strides in understanding cosmology's multiverses, another two levels of Tegmark's hierarchy remain. Level 3 has its origins in quantum theory. Physicists accept that quantum mechanics works brilliantly. It can be used, for example, to calculate a value for the magnetic moment of the electron which matches measurements to 1 part in a billion. However, they have yet to agree on what it means. In the quantum realm, particles don't exist as discrete entities that you can pin down, but as "probability waves". It's the evolution of those

waves that lets quantum physicists predict how electrons cloak an atom, how quarks and gluons interact, and even how objects as large as buckyballs can interfere like light waves.

The pivotal question is what happens to an object's probability wave – its wave function – when someone measures it. Niels Bohr, a founder of quantum mechanics, declared that observing the wave function caused it to collapse and a particle to appear at a particular time and place. That, he said, explains why we see just one outcome out of the infinite possibilities embodied in the wave function.

Yet Bohr's interpretation has long been criticised because it suggests that nothing becomes a reality until someone observes it. In the 1950s, such arguments led Hugh Everett, then a graduate student at Princeton University, to explore what would happen if he jettisoned Bohr's argument. When Everett pictured the wave function rolling majestically on, as the maths of quantum theory said it did, he arrived at a still astonishing and controversial conclusion. Known as the many-worlds interpretation, it calls for the existence of a vast swarm of universes paralleling ours, in which all the possibilities can play out.

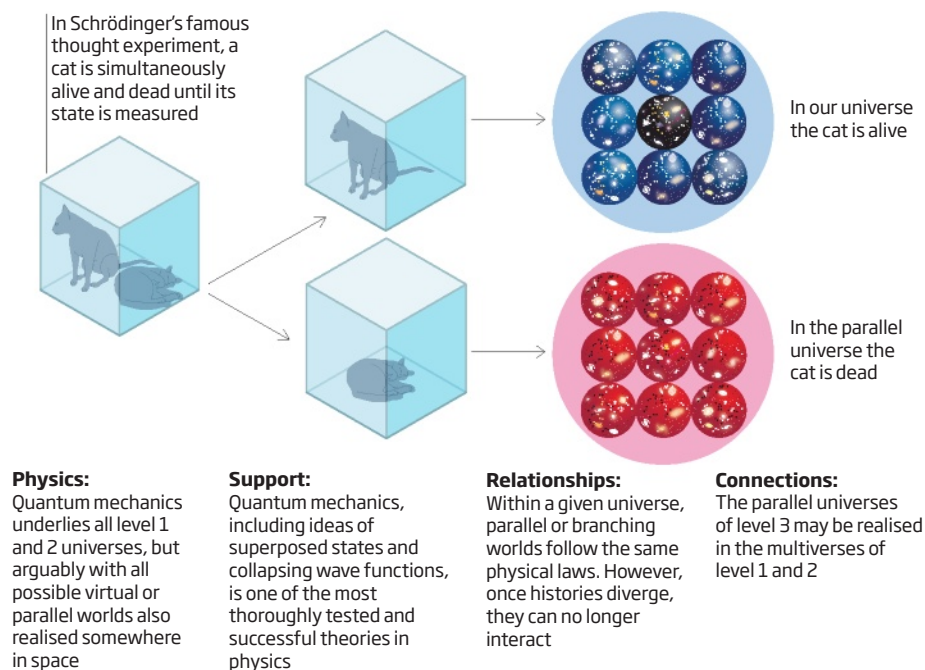
How real are those parallel worlds? As real as dinosaurs, says David Deutsch, a quantum physicist at the University of Oxford. "We've only ever seen fossils, and dinosaurs are the only rational explanation for them," he says. "Many worlds is the only rational explanation for the quantum phenomena we see. I think this is as good as dinosaurs."

The parallel worlds of quantum theory and the multiple universes created by eternal inflation could not seem more different. However, theorists have started to explore the idea that the quantum landscape and the inflationary landscape are one and the same. Bousso and Susskind argue that they produce the same collection of parallel universes. "We both thought for a long time that the multiverse idea and the many-worlds idea were redundant descriptions of the same thing," Susskind says. In 2011, though, he and Bousso developed a consistent way of applying quantum rules to the multiverse. The inflationary multiverse, they conclude, is simply the collection of all the bubble universes that quantum mechanics allows. "The virtual realities of quantum mechanics become genuine realities in the multiverse," Susskind says.

Tegmark also equates the infinite variants of our universe in the level 1 multiverse to the infinity of quantum worlds. "The only

The multiverse hierarchy: level 3

The many-worlds interpretation of quantum mechanics suggests a continually branching series of multiverses



difference between level 1 and level 3," he says, "is where your doppelgängers reside."

The multiverses mentioned so far demote our universe to the status of a pebble in a vast landscape, but at least they allow it to be real. Philosopher Nick Bostrom at the University of Oxford ups the ante by arguing that the universe we experience is just a simulation running on an advanced civilisation's supercomputer.

His argument is simple. Civilisations that last will develop essentially unlimited computing power. Some of those will run "ancestor simulations", reconstructions of their forebears or other beings. Just as millions of us play video games like The Sims, a trans-human civilisation would probably run multiple simulations, making it plausible that we are in one of them. Bostrom doubts that we will find evidence that we do or don't live in a simulation. He argues that an advanced civilisation smart enough to create a simulation in the first place would prevent people inside from noticing a problem, or erase the evidence. Tegmark categorises this and several even more speculative forays as level 4 multiverses.

We might not be able to test or disprove a simulation, but what about other kinds of multiverses? Theorists point to several ways that other universes could have left signs we can see. For example, the bubble universes of eternal inflation can collide. One result could be annihilation. "If you're in a bubble and the wall accelerates towards you, it's bad news," says Matthew Kleban, a theoretical physicist at New York University. But in some cases bubbles bounce apart, leaving a telltale smudge in the CMB. Kleban and his colleagues have calculated the details – a symmetrical spot within a particular range of sizes that stands out from its surroundings because it has a different temperature and polarisation.

Something similar has been seen in the CMB, but Kleban admits that the evidence

for a collision is still weak. Still, he says, a collision could be confirmed by data from the Planck satellite's studies of the CMB, or by future missions. Such a find would provoke a new Copernican revolution. "It would tell us that we're inside a bubble that is embedded with a huge number of other bubbles with different laws of physics," he says. "That would be a very, very important discovery."

Quantum mechanics predicts that emerging universes start out entangled with each other. That early entanglement may leave long-lasting signs. "The moment you have a physical mechanism that shows how the universe is born, you end up with a whole series of predictions of how that universe should look at later times," says Laura Mersini-Houghton, a cosmologist at the University of North Carolina in Chapel Hill. She and her colleagues used this approach to generate four predictions all of which, she says, have been confirmed.

One was the existence of a giant void in our universe. Data from NASA's WMAP satellite and the Sloan Digital Sky Survey show that something similar lurks in the constellation Eridanus. They also predicted that the overall intensity of the CMB should be about 20 per cent less than predicted by inflation and, surprisingly, that everything in our universe is flowing in a particular direction. "That was so outrageous that nobody believed it," Mersini-Houghton says, "but it has been confirmed."

Her group's most controversial prediction involves particle physics. Most physicists expect that collisions at the Large Hadron Collider near Geneva in Switzerland will reveal the first signs of a new symmetry of nature known as supersymmetry. Considered an essential ingredient of string theory, supersymmetry requires that every particle in the standard model has a hefty partner waiting to be discovered. Most researchers believe that the high energies at the LHC are sufficient to create these beasts.

The multiverse hierarchy: level 4

If our universe is just a simulation, there could be infinitely many kinds of universes that differ in arbitrary ways from ours



Physics:
Anything goes

Support:
None. The idea is that long-lived technological civilisations will probably command vast computing power and may choose to run multiple "ancestor simulations" which will soon outnumber natural universes. There may be a one-to-one relationship between mathematics and reality: every conceivable mathematical system may represent a real universe

Relationships:
Arbitrary or non-existent

Connections:
Level 4 must contain the ultimate theory of everything. If so, there's no level 5

However, Mersini-Houghton has predicted that 10,000 times more energy is needed. So far no signs of supersymmetry have appeared. Although some researchers are intrigued, most are taking a wait-and-see attitude towards this ambitious application of quantum mechanics to the multiverse.

Theorists have started to explore other testable predictions that follow from various kinds of multiverses. These include whether space is flat or slightly curved and particles and forces other than those predicted by the standard model or supersymmetry.

Three decades after the concept was born, many researchers now agree that at least some kinds of multiverse stand on firm theoretical foundations, yield predictions that astronomers and particle physicists can test, and help explain why our universe is the way it is. Still, Bousso says, multiverses exist at the frontier of science. "You never know anything for sure when you're working on the edge of knowledge," he says. "But if I didn't think it was by far the best bet, I wouldn't waste my time on it." Susskind agrees, but adds: "We're nowhere near the end of questions. Surprises will happen." ■

"A trans-human civilisation would probably run multiple simulations of the universe, making it plausible we are in one of them"



CHAPTER THREE

THE FABRIC OF THE COSMOS

The search for ripples in space-time could shake up theories a little more than we bargained for, says **Stuart Clark**

Ear on the universe

IT RESEMBLED the Oscars, only with physicists rather than actors. Three hundred of them were gathered in a ballroom in Arcadia, California; another 100 were connected by video link. All of them were waiting for the opening of an envelope.

What the event might have lacked in glamour (sorry, physicists), it made up for in drama. In contrast to the Hollywood awards, the note inside would either make them all winners, or all losers.

The drama had begun six months earlier when scientists around the world had noticed a peculiar signal. They were looking for gravitational waves – ephemeral ripples in the fabric of the universe that are the last untested prediction of Einstein’s general theory of relativity. It is thought that they can be sparked by the collision of stars, the formation of black holes and the great violence of the big bang itself. By the time they travel enormous distances across space and reach Earth, however, the disturbances are only about one-thousandth the width of the smallest atomic nucleus, making them fiendishly difficult to detect. Still, their discovery would revolutionise the way we study the universe.

With that as a goal, it is no wonder a whole generation of gravitational wave researchers have spent their careers developing equipment capable of picking up these ultra-subtle signals. Long on ambition but short on sensitivity, the detectors have seen more than half a century of innovation and improvements to isolate them from the

slightest disturbances. The effort has brought us to the threshold of detection, but the dogged pursuit has come at a price. “An awful lot of people have become conditioned into thinking that no matter what signal we think we see, it is bound to be just noise,” says Martin Hendry, a gravitational wave physicist at the University of Glasgow, UK.

Which is why the upper echelons of the endeavour occasionally “inject” false signals into the data to keep the rank and file on their toes. Although no one could entirely suppress the suspicion that the peculiar signal they had spotted was one of these tests, the researchers set to work with vigour.

After six months of painstaking analysis, everyone was convinced that what they were seeing were gravitational waves from two dead stars colliding to create a black hole. It had come from the direction of the constellation Canis Major, so researchers referred to it as Big Dog.

Having been forbidden from discussing it with the outside world until the analysis was complete, they gathered in California in March 2011 to decide what to do next. “The collaboration took the decision that it was going to go forward with a paper claiming that it had evidence of a detection,” recalls Hendry. He was in the room for the final hurdle: the opening of the envelope to discover whether management had placed the signal in the system. “I remember thinking just how different the world was suddenly going to become if it turned out to be a real signal,” he says.



LEIF PODDHAJSKY

When the envelope – in reality a PowerPoint presentation – was opened, the world remained the same. Knowing what the gravitational waves from a black hole should look like as they ripple across Earth, the management had secretly arranged to shake the detectors in just the right way and at just the right time. Big Dog was an artificial injection.

But one day soon, it could be a real signal.



"So far, the vast majority of our knowledge of the cosmos has come from looking. Now we want to listen to the universe as well"

Gravitational wave detectors on both sides of the Atlantic are currently undergoing major upgrades to heighten their sensitivity. And when the big day comes, it will be as if astronomers have developed a new sense.

So far, the vast majority of our cosmic knowledge has come by looking – using telescopes to collect light and other forms of electromagnetic radiation, including radio, infrared and X-rays. "Now we want to listen to the universe as well," says Stefano Vitale at the University of Trento in Italy.

Catch a wave

Just how do you detect a ripple in space-time? One way is an indirect detection, such as the one recently achieved by the BICEP2 telescope at the South Pole. Its measurements – which await corroboration – detected the imprint of gravitational waves on the cosmic microwave background (CMB), the radiation that filled the universe shortly after the big bang.

The first direct detection, though, may happen at the Virgo interferometer in Cascina, Italy. It consists of two arms at right angles that each stretch for 3 kilometres across the countryside. Laser light travels along the two arms simultaneously, using mirrors at either end to increase the distance travelled to about 100 kilometres.

Ordinarily, the laser beams should take the same time to complete their journeys. This is checked by bringing the beams together at the end of their respective journeys to see if they remain in synchrony with each other. But if a gravitational wave passes through, it will minutely alter the path length of the laser, first in one arm and then in another. This means that the signals will fall out of step and then return to normal.

The same technique is used at the Laser Interferometer Gravitational-Wave Observatory (LIGO) in the US. It has two detectors at Hanford, Washington, and one at Livingston, Louisiana.

Gravitational wave-hunting at the Virgo and LIGO detectors stopped in late 2010, to make way for upgrade work. All are being enhanced to become about 10 times more sensitive. By comparing the signals between the European and American detectors, physicists can triangulate the direction of the incoming gravitational wave and pinpoint its source. ➤

"LISA Pathfinder could become the greatest gravitational experiment since 1919, when Arthur Eddington confirmed Einstein's general theory of relativity"

When operations recommence at the three sites in 2016, the scientists involved believe that it will be only a matter of time before the first detection. "Late 2016 or early 2017, that's about when we expect to have sufficient sensitivity that the odds in favour of detecting something should start to get rather good," says Hendry.

As joyous as the first direct detection will be, the work will be far from over. As the electromagnetic spectrum covers a huge range of wavelengths, from radio waves and microwaves to light and gamma rays, so the gravitational wave spectrum is diverse (see diagram). There are many possible wavelengths given out by different celestial objects. Not all of them are detectable from Earth's surface.

Ground-based detectors can only pick up the higher-frequency waves, which are expected to come from the formation of small black holes caused by either exploding stars or the collisions of dead stellar cores. If we want to detect the gravitational waves caused by supermassive black holes feeding, or merging with others during galaxy collisions, then we need to move our detectors out into space. Up there, the lower-frequency gravitational waves will not be masked by the seismic grumbling of Earth's interior.

In 2015, the European Space Agency (ESA) will launch the LISA Pathfinder spacecraft. Built to test the detection technology, it is the crucial first step needed for a full-blown gravitational wave mission in space.

At the heart of LISA Pathfinder will be two

identical cubes of metal, each made from 2 kilograms of gold and platinum. These metal hearts will not be beating – quite the reverse. "They are going to be the stillest thing in the solar system, and I mean really still. Nothing as still as this has ever been made before," says Vitale, who is principal investigator for the mission. The test masses will be clamped tightly for launch and released once the spacecraft is in orbit. After that, they will float freely inside, separated by a distance of about 35 centimetres.

Modified gravity

An onboard laser mechanism will monitor the two metal cubes for changes in their relative motion. It will detect movement as small as one billionth of a millimetre – a picometre (10^{-12} metres). This is a thousand times more precise than the last such sensor ESA built, which flew on its GOCE mission to measure Earth's gravitational field from orbit.

If all goes well with LISA Pathfinder, the next time such test masses fly, they will be on the full-blown LISA mission with a possible launch date in 2028 or 2034. Then, instead of being just 35 centimetres apart, they will be on three different spacecraft up to 5 million kilometres apart. At this distance, they will be able to use their lasers to detect the slight jostling of the spacecraft as a gravitational wave ripples by.

That would make every gravitational physicist extremely happy, not to mention giving them a lot of work to do in analysing the signals. However, in a twist, there may be more science to LISA Pathfinder than anyone originally imagined.

If a small but growing team of scientists and engineers get their way, LISA Pathfinder could transform itself into more than a technology demonstrator. It could become the greatest gravitational experiment since 1919, when the astrophysicist Arthur Eddington confirmed Einstein's general theory of relativity, the work that extended Isaac Newton's law of gravity to include Einstein's own special theory of relativity.

During a solar eclipse, Eddington showed that light from a distant star cluster was bent on its way to Earth as it passed by the sun. This "lensing" of starlight is something that only happens where gravity is exceptionally strong. It was not anticipated by Newton in his law of gravity, but was predicted by Einstein. Eddington's confirmation was possible because telescopes had become precise enough to measure the effect.

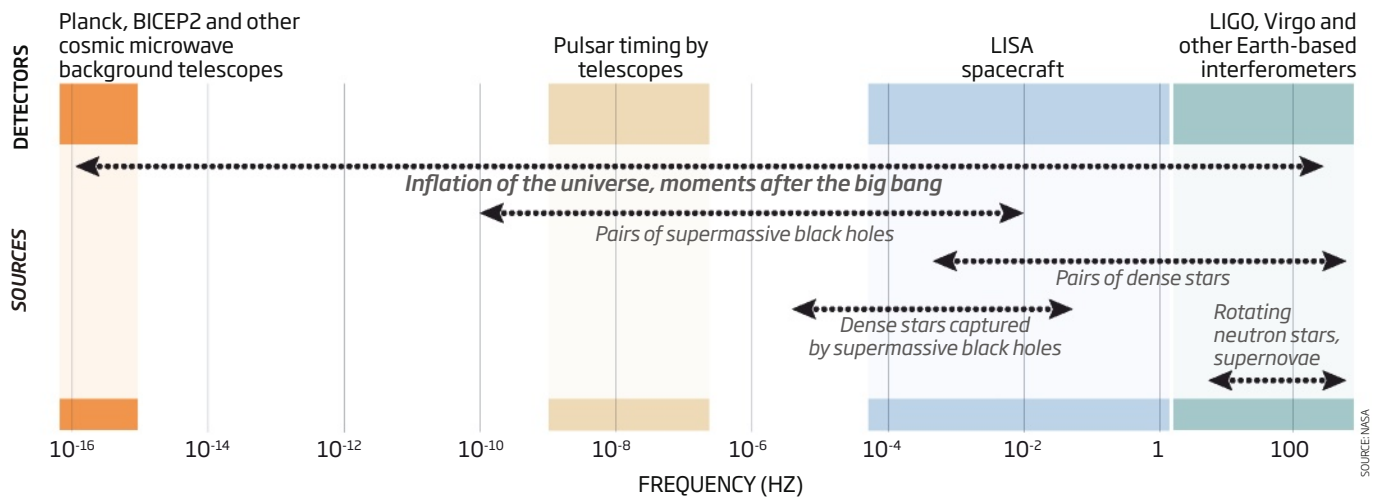
Most physicists assume that Newton's law only breaks down in strong gravity, near a massive celestial body such as a star, galaxy cluster or a black hole. Elsewhere, they assume that gravity behaves as Newton prescribed: dropping by a quarter every time you double

Its mission is to test technology in space, but LISA Pathfinder could do revolutionary science too



Gravity's spectrum

The frequency of gravitational waves depends on the sources they originate from. Detectors focus on different regions of this spectrum



the distance from the centre of a galaxy. This thinking has led to the introduction of dark matter, invisible stuff that provides the gravitational glue keeping individual galaxies rotating, and that binds galaxies to one another in clusters.

Although no one knows what dark matter is, most astronomers believe it exists. Yet many observations can be explained if instead a galaxy has a stronger grip on some stars than Newton predicts.

The most famous of these “modified gravity” scenarios is called modified Newtonian dynamics or MOND. Proposed in the 1980s by Mordechai Milgrom, then at Princeton University, MOND only applies in places where the acceleration due to gravity falls below a certain minuscule value. The locations of such places were pinpointed in 2006 by Jacob Bekenstein of the Hebrew University of Jerusalem in Israel and João Magueijo of Imperial College London. Their calculations revealed certain regions in the solar system, known as saddle points, where the gravity of all the planets, the moons and the sun would cancel out (see “Back in the saddle”, right).

If LISA Pathfinder can be made to pass through or close to one of these saddle points, then its sensitive laser system could test Newton’s law of gravity down into the modified gravity regimes. “If we see something we don’t expect, we can pretty concretely say gravity is not what we thought it was; there’s ‘gravity plus,’” says Ali Mozaffari, also at Imperial College.

Together with Magueijo and others, Mozaffari is investigating a possible extension to the LISA Pathfinder mission. The closer the spacecraft gets to a saddle point, the greater the precision of the test that involves

monitoring the separation of LISA Pathfinder’s metal cubes. If the instrument works as expected and the spacecraft can approach to within 50 kilometres of a saddle point, a MOND-like modification of gravity would stand out like sore thumb. Even with a catastrophic miss by about 400 kilometres, you would still get a strong enough signal to claim a discovery, says Mozaffari.

Since Milgrom’s original proposal, many versions of modified gravity theory have appeared. Whether a signal can be used to pin down a particular one is a question of ongoing

BACK IN THE SADDLE

A gravitational saddle point is the place part way between two celestial objects where their gravitational fields cancel each other out. It is where the balancing point between two objects would be if they were placed on an enormous set of scales.

Saddle points are similar to the Lagrangian points, which are better known because they provide a good location for astronomical instruments. Indeed, the European Space Agency’s LISA Pathfinder mission is destined for a Lagrangian point 1.5 million kilometres away from Earth, towards the sun.

Spacecraft placed at one of these points will feel a force of gravity that keeps them in position relative to Earth as our world orbits the sun. What makes saddle points different is that they are where the gravitational fields cancel out. Pass through a saddle point, and the force of gravity would momentarily drop to zero.

As LISA Pathfinder approaches the Earth-sun saddle point, its payload will detect this drop and see whether it follows Newton’s law of gravity (see main story).

work. “A lot of people have provided predictions about what their theories would say,” says Mozaffari. But to actually tie a signal to a particular theory is harder, he says. “It’s work in progress.” Getting close to a saddle point and seeing nothing unexpected would be easier, because it would allow tight constraints to be placed on modified gravity theories, perhaps even ruling some of them out completely.

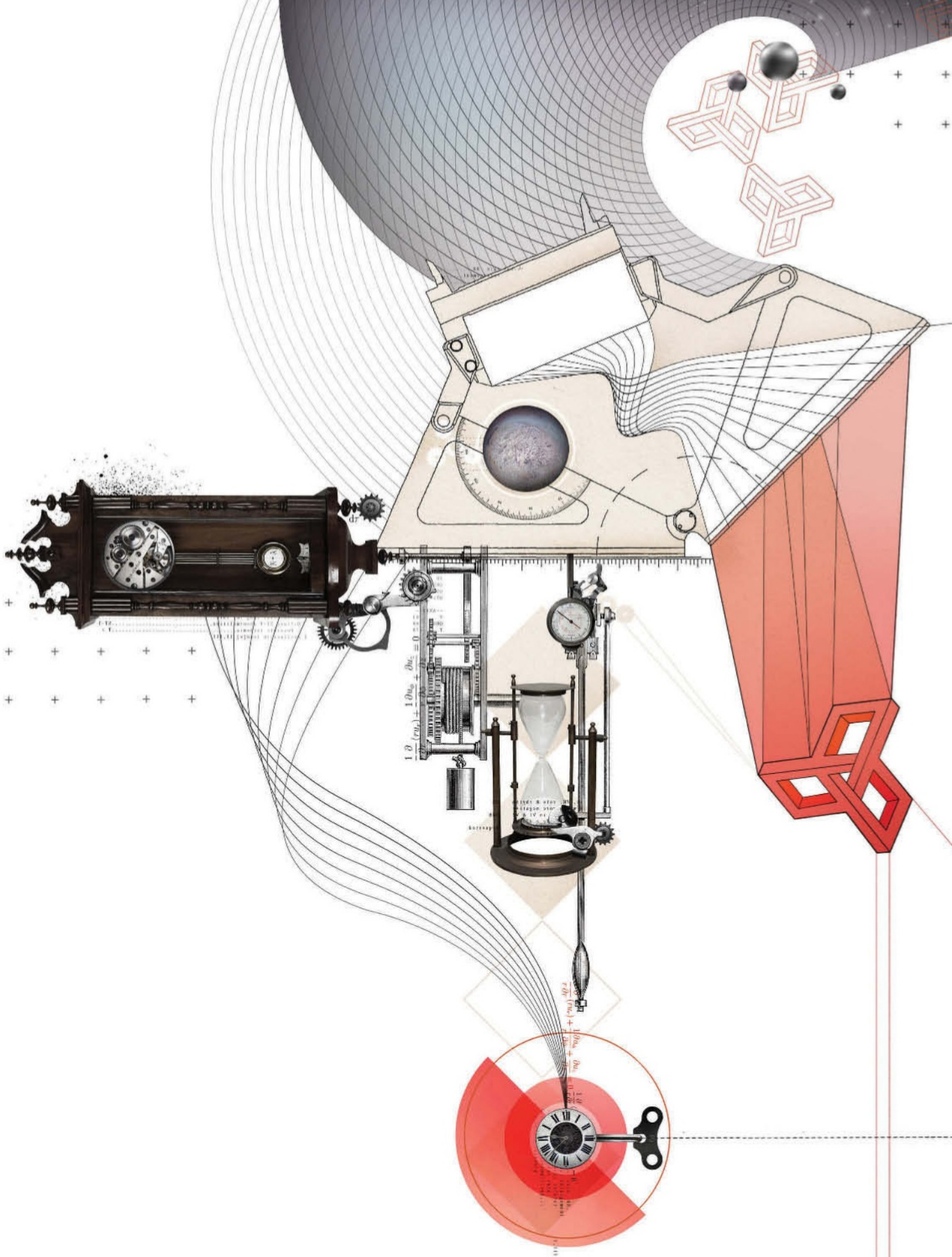
To do this experiment, ESA would have to authorise an extension to the LISA Pathfinder mission for when the gravitational wave technology has been completely demonstrated and understood. Vitale says the demonstration must remain the mission’s top priority. “I wouldn’t say that I am ready to go to the saddle points yet,” he says. “If we have to do extra tests for our gravitational wave detector, that will be our highest priority.”

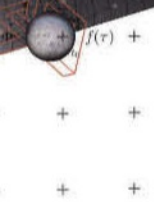
But he does feel the lure of the saddle points. “This is certainly a unique instrument and when you put a unique instrument in a unique place, you are always going to learn something,” says Vitale.

Without doubt, there is a palpable excitement about both LISA Pathfinder and the chances of a gravitational wave detection with a ground-based instrument. “I’m still as enthusiastic as day one,” says Vitale. “You have to be. If you don’t believe, it doesn’t happen.”

As for when that historic first detection will be made, it came up as the final question for debate at a gravitational wave conference in July 2013. The notional date the gathered scientists plumped for was 1 January 2017.

Hendry hopes this is not the case. It may not be the Oscars but, as a Scot, he’s got another party to go to the night before. “I pointed out that I for one would not expect to be anywhere near a computer on that date,” he says. ■





SPACE *against* TIME

Is space the warp and weft of reality,
or time – or both, or neither?
Anil Ananthaswamy picks up the threads

TO ISAAC NEWTON, space was the “sensorium of God”, the organ through which the deity surveyed His creation. It was absolute, unchanging, infinite. Flowing through it “equally without regard to anything external”, as Newton wrote in his great physical treatise *Principia*, was another similarly absolute heavenly creation: time. Not everyone bought that idea. Newton’s perennial antagonist Gottfried Leibniz was notably sniffy of a God who needed an organ to perceive things, and asked pointedly whether a clockmaker deity would need to “wind up his watch from time to time”.

A few centuries on, God features less prominently in the debate, but arguments about the nature of space and time swirl on. Are both basic constituents of reality, or neither – or does one perhaps emerge from the other in some way? We are yet to reach a

conclusive answer, but it is becoming clear that if we wish to make further progress in physics, we must. The route to a truly powerful theory of reality passes through an intimate understanding of space and time.

The search for reality’s building blocks goes to the heart of what physics is about. “When we find the simplest equations for everything in the universe, the fundamental quantities would be what appear in those equations,” says theorist Joe Polchinski of the University of California, Santa Barbara.

That makes it all the more embarrassing that the two sets of equations we use to describe the physical world differ so radically in form and content. Einstein’s relativity, which covers gravity, does a stellar job in describing the universe at large. The theory of quantum mechanics, meanwhile, describes all the other forces and paints a peerless

picture of the world at the smallest scales.

Problems emerge when the large meets the small: in the first instants after the big bang, for example, when the entire universe was a mere pinprick; or at the immense maw of a black hole, whose gravitational pull is so great that not even photons of light can escape. The contradictions that rear up have, at least in part, a very basic origin. “One of the tensions comes from the fact that the relation between space and time is very, very different in general relativity than it is in quantum mechanics,” says theorist Sean Carroll of the California Institute of Technology in Pasadena.

When Einstein developed the theory of special relativity in 1905, it undid Newton’s notions of a clockwork universe, in which objects in an absolute space followed the beat of a heavenly timepiece. Space and time are intertwined into one four-dimensional fabric called space-time. People moving at different speeds measure space-time differently. Just as “here” does not mean the same thing to people in different places, so it is with “now” in Einstein’s relativistic space-time. “What we call now doesn’t have any unique translation to what a person on Alpha Centauri calls now,” says Carroll.

General relativity, which came along in 1916, muddled the waters still further: massive objects curve space-time, and measurements of ruler lengths and clock ticks depend on the strength of the prevailing gravitational field.

In quantum mechanics, things are even more abstruse. A quantum object’s state is described by a wave function, a mathematical object living in an abstract space, known as Hilbert space, that encompasses all the possible states of the object. We can tell how the wave function evolves in time, moving from one state in its Hilbert space to another, using the Schrödinger equation.

In this picture, time is itself not part of the Hilbert space where everything else physical sits, but somehow lives outside it. When we measure the evolution of a quantum state, it is to the beat of an external timepiece of unknown provenance. “Somebody gave us a clock, a grandfather clock, and we just use this clock,” says Nathan Seiberg of Institute for Advanced Study in Princeton, New Jersey.

As for space, its status depends on what you are measuring. The wave function of an electron orbiting the atomic nucleus will include properties of physical space such as the electron’s distance from the nucleus. But the wave function describing the quantum spin of an isolated electron has no mention of space: according to the mathematics, the ➤

picture we often paint of an electron physically rotating is meaningless.

"This is one sense in which there are attributes of physical systems which don't refer to space, but which change in time," says Abhay Ashtekar of Pennsylvania State University. "One could say that for those attributes, time is more fundamental than space."

That is how things stand if you consider general relativity and quantum mechanics in isolation. Relativity says space and time are on the same footing – together they are the fabric of reality. Quantum mechanics, on the other hand, treats time and space differently, with time occasionally seeming more fundamental.

All this starts unravelling, however, whenever we attempt to combine the two theories to find a greater theory capable of describing reality on all scales, large and small.

One of the most ambitious such attempts is string theory. It has an odd relationship with space. One of the theory's key characteristics is the existence of extra dimensions of space so tightly curled up as to be almost undetectable. It needs at least 10 space-time dimensions to be mathematically consistent. But a celebrated result derived in 1997 by theorist Juan Maldacena suggests mathematical trapdoors exist between these different dimensions. According to his "anti-de Sitter/conformal field theory correspondence" – AdS/CFT for short – under certain circumstances, you can swap the fiendishly complex 10D representations of string theory that include gravity for a more tractable 4D representation that dispenses with gravity.

As you do this, one-dimensional time remains seemingly unchanged, but space

is transformed: a point in the 4D world translates to multiple points within the 10D world. "In this example it seems perfectly clear that space is not fundamental. It is very, very different depending on what description of the world you are using," says Carroll.

Black hole drama

But things may not be so clear-cut. Polchinski, for one, has started to wonder if such insights from Maldacena's conjecture are entirely justified. We already know that the AdS/CFT correspondence is only valid for a specific sort of space-time that is not quite the space-time of our universe. The fabric of our universe is geometrically almost flat: two light rays that start out parallel stay so. But the space-time for which the AdS/CFT correspondence is valid is negatively curved, such that two initially parallel light rays slowly start to diverge. The intractable mathematics of string theory means that no one has yet worked out an AdS/CFT-like correspondence for our space-time.

Polchinski's team has raised a more fundamental objection, however – by peering into the dark heart of a black hole. Black holes have a history of testing theories to the limit. Predicted by general relativity and thought to exist where massive stars once lived, as well as at the heart of every galaxy, they have insatiable appetites that appear to gobble up even information, something forbidden by quantum theory. If you try to sidestep the issue by allowing information to escape from a black hole, quantum theory says that a blazing "firewall" of high-energy radiation appears just inside the event horizon, the black hole's point of no return.

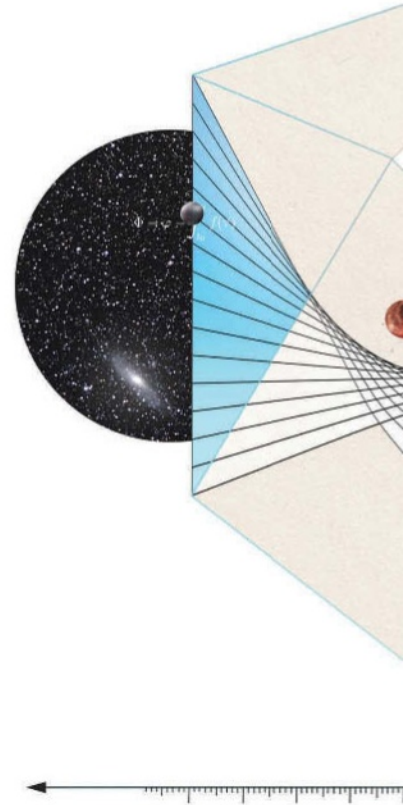
That in turn goes against the predictions of general relativity, which says that anything going past a black hole's event horizon should encounter nothing but gently curved space-time – no theatrics, no drama. If you want to keep quantum mechanics intact and avoid a firewall too, something else about space-time needs to give, such as the dictum that nothing can travel faster than the speed of light. "This does point to the fact that we may be missing something in our conceptual description," says Steve Giddings, also at the University of California, Santa Barbara.

It was to crack this conundrum that Polchinski's team turned to Maldacena's conjecture. They started by chucking a black hole into a volume of negatively curved space-time. If the conjecture were to hold, then the 4D physics of an observer on the surface of the volume should be able to account for the

physics of an observer deep inside the 10D bulk of a black hole held within – only with easier mathematics. Instead, what the two observers see is described by two different quantum theories.

The AdS/CFT correspondence is beloved among string theorists, and Polchinski knows he is going against the grain. "I want to shake people's faith in AdS/CFT," he says. If he succeeds, any conclusions drawn from it and string theory about the status of space and time may not be the last word.

But string theory is just one as-yet-unproven approach to unifying relativity and quantum theory. Another, known as loop quantum gravity, has its origins in the mid-1980s when Ashtekar rewrote Einstein's equations of general relativity using a quantum mechanical framework. Working with physicists Lee Smolin and Carlo Rovelli, he used these equations to arrive at a picture in which space-time is a smooth fabric – until, as with any fabric, you look at it very closely. In the case of the space-time of loop quantum gravity, if you zoom in to tiny scales of about 10^{-35} metres, a distance known as the Planck length, you see a warp and weft of loops of gravitational field lines.



"Relativity says space and time are on the same footing. Quantum mechanics treats them differently, with time more fundamental"



"A light ray always moves at one unit of space per unit of time – in a sense, it is on the edge between space and time"

As with string theory, the equations of loop quantum gravity have proved difficult to work with, and have produced little in terms of verifiable experimental predictions. Still, they provide a different perspective on the primacy of space and time. Chunks of space, one Planck length to a side, appear first in the theory, while time pops up only later as an expression of the relationships between other observable physical properties. For instance, you can define the tick of a clock in terms of changes in the gravitational field and then observe how another field, say the electromagnetic field, changes with respect to the "ticking" of the gravitational field. Here, both space and time seem to emerge from something deeper, says Ashtekar. "But somehow space might emerge first, and time is born by observing relations between various physical subsystems."

In a sense, that's coming full circle to how Newton viewed some aspects of time. Although seeing absolute time as something God-given, he recognised that we measure a "common" time relationally – that is, by keeping track of other properties. Earth's motion around the sun in space represents a unit of time, for example – the year – which we subdivide or multiply to calculate the duration

of anything else, like the length of a season.

We do not know as yet what the deeper thing might be from which space and time emerge, but Ashtekar is not the only one thinking along these lines. Seiberg has a similar intuition based on findings in string theory, which suggest space is emergent, and general relativity, which says that space and time are interwoven. "We have many examples where we have emergent space and, given that space and time mix, there's no doubt that time will also be emergent," he says.

Relaxing restrictions

Giddings is also exploring the idea that neither space nor time is fundamental, using that most uncomfortable of theoretical test beds. He has been trying to describe a black hole, from its interior to way outside its event horizon, using a network of interconnected quantum-mechanical Hilbert spaces that do not presuppose the existence of space or time. This allowed him, for example, to relax the restriction that nothing can travel through space-time faster than light – imposed by Einstein's relativity – and see what would happen as a result.

By observing how one attribute of a quantum system changes with respect to another in this set-up, Giddings showed in 2012 how time can emerge relationally, in another nod to Newton's "common" time. A concept of space also emerges from his calculations by loosely specifying different Hilbert spaces to correspond to different physical locations, such as the interior of a black hole.

Ultimately, Giddings thinks that even this notion of space can be linked to the dynamics of the system, just like time. In this case, neither would be primary.

It is early days for the idea. To be taken seriously, Giddings must show that general relativity and the normal picture of space-time can be derived from his network of Hilbert spaces for situations that are not as extreme as black holes. "We have a way to go to fully understand this," he says.

Polchinski is also tugging at that constant speed of light in relativity. In a sense, this provides a reference of both space and time. A light ray always moves at one unit of space per unit of time – a constant diagonal on any graph of space against time.

"The direction that light rays travel in is neither space nor time; we call it 'null'. It's on the edge between space and time," says Polchinski. "A lot of people have this intuition that in some sense the existence of these null directions might be more fundamental than space or time."

Will that or any other intuition lead us anywhere in our quest towards a greater theory? The current impasse in physics is not unlike the situation in the early 20th century, before Einstein's master stroke replaced space and time with one space-time. Today we appear to be facing a similar obstruction to progress. Many potential ways around lead to different worlds of space and time – and we have as yet little clue which route to follow. Space, time, both or neither? Perhaps only time will tell. ■



Does some deeper level of reality lurk beneath Einstein's universe? **Amanda Gefter** investigates

Beyond space-time

IT WASN'T so long ago we thought space and time were the absolute and unchanging scaffolding of the universe. Then along came Albert Einstein, who showed that different observers can disagree about the length of objects and the timing of events. His theory of relativity unified space and time into a single entity—space-time. It meant the way we thought about the fabric of reality would never be the same again. “Henceforth space by itself, and time by itself, are doomed to fade into mere shadows,” declared mathematician Hermann Minkowski. “Only a kind of union of the two will preserve an independent reality.”

But did Einstein's revolution go far enough? Physicist Lee Smolin at the Perimeter Institute for Theoretical Physics in Waterloo, Ontario,



Momentum space isn't as alien as it first sounds. When you look at the world around you, says Smolin, you don't ever observe space or time – instead you see energy and momentum. When you look at your watch, for example, photons bounce off a surface and land on your retina. By detecting the energy and momentum of the photons, your brain reconstructs events in space and time.

The same is true of physics experiments. Inside particle smashers, physicists measure the energy and momentum of particles as they speed toward one another and collide, and the energy and momentum of the debris that comes flying out. Likewise, telescopes measure the energy and momentum of photons streaming in from the far reaches of the universe. “If you go by what we observe, we don't live in space-time,” Smolin says. “We live in momentum space.”

And just as space-time can be pictured as a coordinate system with time on one axis and space – its three dimensions condensed to one – on the other axis, the same is true of momentum space. In this case energy is on one axis and momentum – which, like space, has three components – is on the other (see diagram, page 48).

Simple mathematical transformations exist to translate measurements in this momentum space into measurements in space-time, and the common wisdom is that momentum space is a mere mathematical tool. After all, Einstein showed that space-time is reality's true arena, in which the dramas of the cosmos are played out.

Smolin and his colleagues aren't the first to wonder whether that is the full story. As far back as 1938, the German physicist Max Born noticed that several pivotal equations in quantum mechanics remain the same whether expressed in space-time coordinates or in momentum-space coordinates. He wondered whether it might be possible to use this connection to unite the seemingly incompatible theories of general relativity, which deals with space-time, and quantum mechanics, whose particles have momentum and energy. Maybe it could provide the key to the long-sought theory of quantum gravity.

Born's idea that space-time and momentum space should be interchangeable – a theory now known as “Born reciprocity” – had a remarkable consequence: if space-time can be curved by the masses of stars and galaxies, as Einstein's theory showed, then it should be possible to curve momentum space too.

At the time it was not clear what kind of physical entity might curve momentum

Canada, doesn't think so. He and a trio of colleagues are aiming to take relativity to a whole new level, and they have space-time in their sights. They say we need to forget about the home Einstein invented for us: we live instead in a place called phase space.

If this radical claim is true, it could solve a troubling paradox about black holes that has stumped physicists for decades. What's more, it could set them on the path towards their heart's desire: a “theory of everything” that will finally unite general relativity and quantum mechanics.

So what is phase space? It is a curious eight-dimensional world that merges our familiar four dimensions of space and time and a four-dimensional world called momentum space.

“If Einstein's space-time is no longer something all observers can agree on, is it the true fabric of reality?”

LUKE BROOKES

space, and the mathematics necessary to make such an idea work hadn't even been invented. So Born never fulfilled his dream of putting space-time and momentum space on an equal footing.

That is where Smolin and his colleagues enter the story. Together with Laurent Freidel, also at the Perimeter Institute, Jerzy Kowalski-Glikman at the University of Wroclaw, Poland, and Giovanni Amelino-Camelia at Sapienza University of Rome in Italy, Smolin has been investigating the effects of a curvature of momentum space.

"Relative locality deals a huge blow to our understanding of the nature of reality"

The quartet took the standard mathematical rules for translating between momentum space and space-time and applied them to a curved momentum space. What they discovered is shocking: observers living in a curved momentum space will no longer agree on measurements made in a unified space-time. That goes entirely against the grain of Einstein's relativity. He had shown that while space and time were relative, space-time was the same for everyone. For observers in a curved momentum space, however, even space-time is relative (see diagram, page 49).

This mismatch between one observer's space-time measurements and another's grows with distance or over time, which means that although space-time in your immediate vicinity will always be sharply defined, objects and events in the far distance

become fuzzier. "The further away you are and the more energy is involved, the larger the event seems to spread out in space-time," says Smolin.

For instance, if you are 10 billion light years from a supernova and the energy of its light is about 10 gigaelectronvolts, then your measurement of its location in space-time would differ from a local observer's by a light second. That may not sound like much, but it amounts to 300,000 kilometres. Neither of you would be wrong – it's just that locations in space-time are relative, a phenomenon the researchers have dubbed "relative locality".

Relative locality would deal a huge blow to our picture of reality. If space-time is no longer an invariant backdrop of the universe on which all observers can agree, in what sense can it be considered the true fabric of reality?

That is a question still to be wrestled with, but relative locality has its benefits, too. For one thing, it could shed light on a stubborn puzzle known as the black hole information-loss paradox. In the 1970s, Stephen Hawking discovered that black holes radiate away their mass, eventually evaporating and disappearing altogether. That posed an intriguing question: what happens to all the stuff that fell into the black hole in the first place?

Relativity prevents anything that falls into a black hole from escaping, because it would have to travel faster than light to do so – a cosmic speed limit that is strictly enforced. But quantum mechanics enforces its own strict law: things, or more precisely the information that they contain, cannot simply vanish from reality. Black hole evaporation put physicists between a rock and a hard place.

According to Smolin, relative locality saves

Just where in space-time did the Cygnus loop supernova remnant form?



J.J. HESTER (ARIZONA STATE UNIVERSITY)/NASA

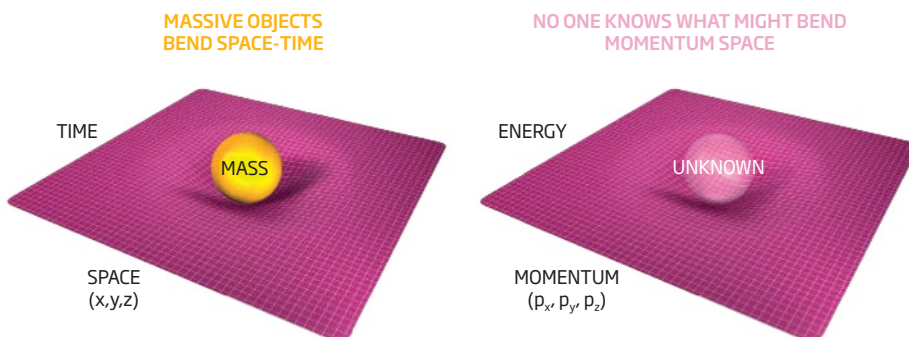
the day. Let's say you were patient enough to wait around while a black hole evaporated, a process that could take billions of years. Once it had vanished, you could ask what happened to, say, an elephant that had once succumbed to its gravitational grip. But as you look back to the time at which you thought the elephant had fallen in, you would find that locations in space-time had grown so fuzzy and uncertain that there would be no way to tell whether the elephant actually fell into the black hole or narrowly missed it. The information-loss paradox dissolves.

Big questions still remain. For instance, how can we know if momentum space is really curved? To find the answer, the team has proposed several experiments.

One idea is to look at light arriving at the Earth from distant gamma-ray bursts. If momentum space is curved in a particular way that mathematicians refer to as "non-metric", then a high-energy photon in the gamma-ray burst should arrive at our telescope a little later than a lower-energy photon from the same burst, despite the

Fabrics of reality

Space-time is like a malleable sheet with the three spatial coordinates on one side and time on the other. Momentum space is similar, with three coordinates of momentum and energy





energy photons, which would account for the observed delays.

In order to disentangle the properties of the explosions from properties of relative locality, we need a large sample of gamma-ray bursts taking place at various known distances. If the delay is a property of the explosion, its length will not depend on how far away the burst is from our telescope; if it is a sign of relative locality, it will. Amelino-Camelia and the rest of Smolin's team are watching carefully as more data come in.

The questions don't end there, however. Even if Fermi's observations confirm that momentum space is curved, they still won't tell us what is doing the curving. In general relativity, it is momentum and energy in the form of mass that warp space-time. In a world in which momentum space is fundamental, could space and time somehow be responsible for curving momentum space?

Work by Shahn Majid, a mathematical physicist at Queen Mary University of London, might hold some clues. In the 1990s, he showed that curved momentum space is equivalent to what's known as a noncommutative space-time. In familiar space-time, coordinates commute – that is, if we want to reach the point with coordinates (x, y) , it doesn't matter whether we take x steps to the right and then y steps forward, or if we travel y steps forward followed by x steps to the right. But mathematicians can construct space-times in which this order no longer holds, leaving space-time with an inherent fuzziness.

In a sense, such fuzziness is exactly what you might expect once quantum effects take hold. What makes quantum mechanics different from ordinary mechanics is Heisenberg's

uncertainty principle: when you fix a particle's momentum – by measuring it, for example – then its position becomes completely uncertain, and vice versa. The order in which you measure position and momentum determines their values; in other words, these properties do not commute. This, Majid says, implies that curved momentum space is just quantum space-time in another guise.

What's more, Majid suspects that this relationship between curvature and quantum uncertainty works two ways: the curvature of space-time – a manifestation of gravity in

“There would be no way to tell whether an elephant actually fell into the black hole or narrowly missed it”

Einstein's relativity – implies that momentum space is also quantum. Smolin and colleagues' model does not yet include gravity, but once it does, Majid says, observers will not agree on measurements in momentum space either. So if both space-time and momentum space are relative, where does objective reality lie? What is the true fabric of reality?

Smolin's hunch is that we will find ourselves in a place where space-time and momentum space meet: an eight-dimensional phase space that represents all possible values of position, time, energy and momentum. In relativity, what one observer views as space, another views as time and vice versa, because ultimately they are two sides of a single coin – a unified space-time. Likewise, in Smolin's picture of quantum gravity, what one observer sees as space-time another sees as momentum space, and the two are unified in a higher-dimensional phase space that is absolute and invariant to all observers. With relativity bumped up another level, it will be goodbye to both space-time and momentum space, and hello phase space.

“It has been obvious for a long time that the separation between space-time and energy-momentum is misleading when dealing with quantum gravity,” says physicist João Magueijo of Imperial College London. In ordinary physics, it is easy enough to treat space-time and momentum space as separate things, he explains, “but quantum gravity may require their complete entanglement”. Once we figure out how the puzzle pieces of space-time and momentum space fit together, Born's dream will finally be realised and the true scaffolding of reality will be revealed. ■

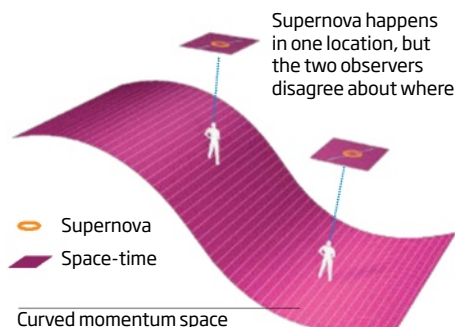
two being emitted at the same time.

Just that phenomenon has already been seen, starting with some unusual observations made by a telescope in the Canary Islands in 2005. The effect has since been confirmed by NASA's Fermi gamma-ray space telescope, which has been collecting light from cosmic explosions since it launched in 2008. “The Fermi data show that it is an undeniable experimental fact that there is a correlation between arrival time and energy – high-energy photons arrive later than low-energy photons,” says Amelino-Camelia.

Still, he is not popping the champagne just yet. It is not clear whether the observed delays are true signatures of curved momentum space, or whether they are down to “unknown properties of the explosions themselves”, as Amelino-Camelia puts it. Calculations of gamma-ray bursts idealise the explosions as instantaneous, but in reality they last for several seconds. Although there is no obvious reason to think so, it is possible that the bursts occur in such a way that they emit lower-energy photons a second or two before higher-

Where's the supernova?

Observers in curved momentum space disagree about where in space-time an event such as a supernova takes place







Strange signals travelling from distant galaxies hint at turbulence for Einstein's theory of space-time, says Stuart Clark

Warning light

WE LIVE in an invisible landscape: a landscape that, although we cannot perceive it directly, determines everything that we see and do. Every object there is, from a planet orbiting the sun to a rocket coasting to the moon or a pencil dropped carelessly on the floor, follows its imperceptible contours. We battle against them each time we labour up a hill or staircase.

This is the landscape of space-time: the underlying fabric of the physical universe, perhaps of reality itself. Although we don't see its ups and downs, we feel them as the force

"Space-time is the fabric of the universe, perhaps of reality itself. But no one knows what it is"

we call gravity. Developed by the physicist Hermann Minkowski in the 20th century, and used by Albert Einstein in his general theory of relativity, space-time has become one of the most powerful concepts in all of physics.

There is just one nagging problem: no one knows what it is. Einstein envisaged space-time as a perfectly smooth surface warped by the mass of stars, planets and galaxies to produce gravity. But signals from a variety of celestial objects are hinting at something different. If the observations are confirmed – and they are controversial – they suggest that the landscape of reality is altogether more rugged than Einstein thought. That would mean his isn't the last word on space-time or gravity, and would change our perception of the universe fundamentally.

Before Einstein, space and time were thought to be separate properties of the universe. For Isaac Newton, they were a rigid framework of creation, and perhaps even some sort of embodiment of God – a "sensorium" through which He viewed the world – with gravity and movement the Almighty's will made manifest. For many, this strayed too far into the realms of maverick theology, and Newton's religious interpretations were soon sidelined. But few questioned the underlying science.

Only in the mid-19th century did it become clear that Newton's dynamics couldn't explain the subtleties of Mercury's orbital motion around the sun. Einstein's relativity could, but only by melding space and time into one mathematically indistinguishable whole, in which what happened to one also affected the other: the space-time continuum.

But although the mathematics of relativity describes space-time's properties very well, it is silent on its underlying nature. We are left to scratch around for observational clues. Everything in the universe, from the largest galaxy to the smallest particle, the dullest radio wave to the brightest ray of light, is immersed in space-time and so presumably must interact with it in some way. The question becomes whether those interactions imprint any signature that we might measure and interpret, and so see the true physical guise of space-time.

"This is a beautiful question, and we are at the beginning of answering it," says Giovanni Amelino-Camelia of the La Sapienza University of Rome in Italy.

In 2005, we seemed to have glimpsed an answer. MAGIC – the Major Atmospheric ➤

Gamma-ray Imaging Cherenkov telescope – is a series of giant receivers on La Palma in Spain’s Canary Islands tuned to detect cosmic light of the highest energy: gamma rays. On the night of 30 June, the array detected a burst of gamma radiation from a giant black hole at the heart of Markarian 501, a galaxy some 500 million light years away. This wasn’t so unexpected. Our theories predict that every time something falls into such a black hole, a flare of radiation will be given off. But those large enough to be caught by an earthbound telescope, even a mighty receiver like MAGIC, are few and far between, and the Markarian flare was pretty much the first of its type to be seen.

Quantum foam

And detailed analysis revealed something decidedly unusual about the burst: the lower-energy radiation seemed to have arrived up to 4 minutes before the higher-energy radiation. This is a big no-no if space-time behaves according to Einstein’s relativity. In relativity’s smooth space-time, all light travels at the same speed regardless of its energy. But the effect was entirely compatible with other, rival theories that attempt to characterise space-time in terms of quantum mechanics – the theory entirely separate to, and incompatible with, general relativity that explains how everything besides gravity works.

In quantum theory, nothing is static or certain. Particles and energy can fluctuate and pop in and out of existence on the briefest of timescales. Many theories of quantum gravity – the yearned-for “theories of everything” that will unify our descriptions of space-time and gravity with quantum mechanics – suggest something similar is true of space-time: instead of a smooth continuum, it is a turbulent quantum foam with no clearly defined surface. Einstein’s undulating landscape becomes more like a choppy seascape through which particles and radiation must fight their way. Lower-energy light with its longer wavelengths would be akin to an ocean liner, gliding through the foamy quantum sea largely undisturbed. Light of higher energy and shorter wavelengths, on the other hand, would be more like a small dinghy battling through the waves.

In 1998, Amelino-Camelia and John Ellis, then at CERN near Geneva, Switzerland, had proposed that high-energy light from distant, active galaxies could be used to check for this effect. The huge distance would allow for even subtle effects to build into a detectable time

lag. On the face of it, this was exactly what MAGIC had seen.

Things are seldom that simple in physics, and the MAGIC observations have generated lively discussion. “This has become quite a musical,” says Robert Wagner of the Max Planck Institute for Physics in Munich, Germany, part of the team that made the initial observation. When a similar gamma-ray telescope, HESS – the High Energy Stereoscopic System in the Namibian outback – caught sight of another giant galactic flare in July 2006, it was the perfect opportunity to test the theory. The galaxy in question, PKS 2155-304, is four times as far away from Earth as Markarian 501, so the time delay should have been even bigger.

But... nothing. “We saw no hint of a time delay,” says Agnieszka Jacholkowska, one of the team analysing the signals at Pierre and Marie Curie University in Paris, France. If we assume that space-time, whatever it may be, is probably the same everywhere, this suggests that the original time delay was something intrinsic to the source of the gamma rays in Markarian 501. It is conceivable, for example, that particles were accelerated along magnetic fields near the centre of the galaxy, which would naturally result in the emission of lower-energy gamma rays first. But since no one quite knows what processes take place in these dark galactic hearts, there was still plenty of room for debate.

And so things remained until 2013, when the most energetic gamma rays ever seen in our short history of observations hit Earth.

It was a gamma-ray burst (GRB): a short, intense flash of radiation not from the heart of an active galaxy but from the explosive

“Nothing is static or certain. Particles and energy can pop in and out of existence on the briefest timescales”

death of a hypergiant star. GRBs are so bright that modern telescopes can see them across the entire universe, meaning that their light has travelled through space-time for several billions of years.

Even so, the one observed by NASA’s Fermi telescope on 27 April 2013 – known prosaically as GRB130427A – was eye-popping. It showered Earth with 10 times as many high-energy gamma rays as a run-of-the-mill burst, and included one gamma-ray photon that carried 35 billion times more energy than a visible



ALEX CHERNEY, TERRASTRO.COM/SPL

photon. Automatic alerts were sent out to observatories across the world and within hours a battery of telescopes was scrutinising the burst’s aftermath. One of the scientists alerted was Amelino-Camelia.

The following month, he and his colleagues circulated a paper claiming to see a time lag of hundreds of seconds between the burst’s lower- and higher-energy gamma rays. “The numbers work out remarkably well. This is the first time there is robust evidence of this feature,” says Amelino-Camelia.

Robust, because unlike the Markarian 501 observations, it was possible to match the arrival times of photons of various energies with those predicted by a simple equation. This relationship is pleasing to the mathematical eye and might also help us to see what lies beyond relativity if it is indeed broken: different variants of quantum gravity sketch different pictures of space-time and might have different effects on light.

In string theory, for instance, quantum space-time is a tangle of six extra dimensions of space, in addition to the usual three of space and one of time. Photons of different energies will propagate through this arrangement in quite a different way than is predicted in loop quantum gravity, another popular theory that imagines space-time as a form of chain mail composed of interwoven loops.

For the time being, Amelino-Camelia has banned his team from investigating which,



Has MAGIC seen signs of quantum space-time?

if any, of these competing theories is closest to the mark. “For the moment, I think it’s important to keep separate how we wish nature was in theory and how nature really is from the facts we have,” he says.

Instead, the next stage is to see what predictions the time-delay equation makes about time lags in bursts of radiation from other sources. In their paper, Amelino-Camelia and his team report four other GRBs whose behaviour was consistent with the equation, although not conclusively in support. Others find no such evidence. Just days after Amelino-Camelia’s paper came out, Jacholkowska and her colleagues published their analysis of four other, less energetic GRBs observed by the Fermi telescope. They found no hint of time lags.

In Jacholkowska’s view we cannot draw any firm conclusions, because Amelino-Camelia’s interpretation assumes, like the Markarian 501 analysis before it, that the gamma rays were emitted simultaneously regardless of their energy. This is always going to be a problem as long as interpretations are based on single observations of one type of source, Ellis says. “If you found an effect that was similar in two, you’d really begin to think you had found something,” he says.

One test that might clear things up involves neutrinos. These ghostly particles travel at virtually the speed of light, interacting with hardly anything. Because they carry energy,

however, they should interact with space-time, and, if Amelino-Camelia is correct, suffer an energy-dependent time lag – although one that is only measurable if we can find neutrinos that have travelled far enough.

That was always a problem. Nuclear fusion reactions make the sun such a prodigious neutrino factory that it washes out almost all signals from further away. Besides solar neutrinos, the only cosmic neutrinos ever seen have been from the supernova SN1987A, a star that just happened to explode in our cosmic backyard, in the Large Magellanic

“In string theory, quantum space-time is a tangle of six extra dimensions of space on top of the usual three”

Cloud some 170,000 light years away. This is still too close for its neutrinos to manifest any measurable time lag.

Decisive help could now be at hand. IceCube is a neutrino detector buried in a cubic kilometre of Antarctic ice that came fully on stream in 2011. In April 2012, it found two neutrinos that set tongues wagging. Called, in a fit of whimsy, Bert and Ernie, after two characters from the TV show *Sesame Street*, they were far more energetic than those generated by the sun. For that reason alone,

Dan Hooper of Fermilab in Batavia, Illinois, thinks it’s likely that they come from a gamma-ray burst. “There aren’t that many things that can make that amount of energy in a single particle. GRBs top the list,” he says. In May 2013, IceCube announced the discovery of a further 26 neutrinos whose energies possibly betrayed an extragalactic source.

Amelino-Camelia thinks he has found three more in earlier IceCube data – ones that perfectly fit the idea of quantum space-time effects taking place. They all arrived from the general direction of three independently verified GRBs – but, if they are indeed associated with the bursts, got to Earth thousands of seconds earlier than the gamma rays.

Neutrinos are expected to escape from a collapsing star sooner than the light of a GRB because they don’t interact, whereas the visible blast has to fight its way through the collapsing gas before speeding through space. But even taking this into account, Amelino-Camelia maintains that the huge size of the gap between the neutrinos and gamma-ray light is consistent with the different effects of a space-time interaction on them.

Ellis remains sceptical. “Every once in a while, somebody gets a little bit excited but I don’t think there’s any statistically solid evidence yet,” he says. “One of the problems is that extraordinary claims require extraordinary proof, so you have to do something that is really convincing.”

That will inevitably require larger telescopes capable of spotting more gamma rays and neutrinos more quickly. Wagner is involved in an international collaboration of more than 1000 researchers from 23 countries that is aiming to build a giant successor to MAGIC and HESS. The Cherenkov Telescope Array would be 10 times as sensitive, and capable of seeing between 10 and 20 active galaxy flare-ups every year. After spending years developing technology and looking at possible locations, with funding so far mainly from the governments of Germany, Spain and the UK, the collaboration will now be looking for the €200 million needed to turn the telescope into a reality.

Will it finally open our eyes to the landscape around us? Those involved hope so. “There is no reason to be pessimistic,” says Wagner. To find any kind of structure in space-time would be a revolution to rival Einstein’s, and could show the way forward when physics is struggling to see its next step. “It would be hard to overstate how important that would be,” says Hooper. ■

THE FORGOTTEN FORCE

It's not just gravity that shapes the universe.
Katia Moskvitch is attracted to a cosmic mystery

WHEN it comes to the cosmos, gravity is the big attraction. The same force that keeps our feet on the ground also shapes the universe. It takes clouds of gas and sculpts them into planets and stars. It fashions hundreds of billions of stars into galaxies, which clump together to form clusters, then superclusters. Yet gravity isn't the only player in the game – another force operates across the cosmic landscape, and that is magnetism.

Magnetic fields stretch for vast distances in the near-nothingness of deep space, even spanning the billions of light years between galaxies. Admittedly, these fields are feeble. A fridge magnet is more than a million times stronger than the weak, all-pervading sea of magnetism in the Milky Way and beyond. That might explain why cosmology has largely ignored magnetism. After all, how could something so puny influence a galaxy?

Times and minds are changing, however. Yes, gravity holds things together, but key physical processes in the universe need magnetism – from star formation to black holes pumping out high-energy jets. “It turns out that many previously unsolved problems in astronomy suddenly make sense once one includes the effects of interstellar magnetism,”

says Bryan Gaensler of the University of Sydney in Australia.

Could the same be true for the universe at large? What makes fields on the galactic scale and beyond so enticing is that they appear to be the legacy of processes that happened shortly after the big bang. What's more, most of the visible mass in the universe consists of charged particles whose movements are in thrall to cosmic magnetism as well as gravity. That raises the tantalising possibility that magnetism has played a key role in shaping the universe ever since the start of time.

But before we can be certain, we need answers to some important questions: how exactly did the fields get there and when?

We've always known magnetism plays an important part for us close to home. It was German physicist Carl Friedrich Gauss who, in 1835, first measured Earth's magnetic field with the help of a simple magnet suspended on a string. Now we have a pretty good idea how the sun and Earth generate their fields. As molten iron in Earth's outer core (or plasma in the sun's case) moves across magnetic field lines, the effect is to induce electric currents. These give rise to a magnetic field that supplements the existing one. Thanks to this dynamo action, a small “seed”



SIMON DANAHER



field can grow into a much larger one.

And a good job too. Earth's magnetic field shields us from deadly particles that would otherwise tear away the ozone layer and expose our planet to harmful ultraviolet rays. The sun's field protects us too, deflecting yet more deadly particles from beyond the solar system. On a bigger scale, magnetism might even have contributed to the origin of life (see "Power of the weak", page 57).

Few people expected interstellar space to be magnetic, though. The first proof came in 1949, when US astronomers John Hall and William Hiltner showed that "something" was polarising starlight on its way to us. That something turned out to be cosmic magnetism, lining up interstellar dust grains like tiny compass needles. It was a remarkable discovery, says Gaensler.

Since then, a host of techniques have been developed to measure magnetic fields in the Milky Way, its neighbouring galaxies and beyond. For example, a detailed map of the Milky Way's magnetism was made in 2011 by Niels Oppermann, then at the Max Planck Institute for Astrophysics in Garching, Germany, and colleagues. It shows that the field lines follow the galaxy's spiral shape. It also confirms that the Milky Way's overall field is a few microgauss (10^{-6} gauss) – some 100,000 times smaller than the magnetic field at Earth's surface.

Astronomers believe that the magnetic fields in spiral galaxies like the Milky Way are amplified and maintained by a dynamo. As the galaxy turns, it moves charged particles across an existing magnetic field, which boosts the field further. "The hypothesis is that the galaxy is born with a much weaker field," says Larry Widrow, an astrophysicist at Queens University in Kingston, Ontario, Canada. "But where did the weak primordial fields that seed the dynamo come from?" he asks.

The first field

For decades, researchers have been trying to solve this conundrum, but their instruments have not been sensitive enough to test any of their theories.

And so the theories have multiplied. One idea is that the first magnetic fields were produced in very early stars and then spread into the interstellar medium via stellar winds or supernova explosions. Another is that when the first galaxies formed, roughly 100 million years after the big bang, the supermassive black holes at their centres generated strong magnetic fields that were ➤

subsequently dispersed into the intergalactic medium by powerful jets. A more recent idea is that cosmic magnetism may have been triggered by plasma fluctuations in young galaxies. Once you have a weak field, it can be rapidly amplified by the dynamo effect. So the rotation and turbulence of the galaxy's interstellar medium (the gas and dust between stars) reinforce the original, feeble field. These processes can double the strength of a field in roughly the time it takes a star or the centre of a black hole to rotate. Such timescales are negligible compared to the age of the universe, and so a meagre field can quickly reach appreciable strengths.

The trouble is that young galaxies whose light comes to us from far away should have much smaller fields than nearby ones. But astronomers find mounting evidence that microgauss fields existed in these early galaxies too. Either the dynamo mechanism is stronger, or more tantalisingly, the seed fields were generated within a blink of the big bang.

Boom time

Widrow and his colleague Michael Turner at the University of Chicago suggested just such a scenario in 1988. They proposed that primordial magnetic fields were created in the first fleeting moments after the big bang and amplified by inflation, when the budding cosmos expanded faster than light. The large galaxy structures that we observe today were caused by quantum fluctuations in energy during this period. Widrow and Turner showed that inflation could also amplify fluctuations in the electromagnetic field in such a way as to leave the entire universe with a dose of magnetism.

For that idea to work, the pair had to make changes to the famous equations formulated by James Clerk Maxwell to describe electrical and magnetic fields, and they introduced an exotic particle called the axion. "The idea was ad hoc and aspects of the proposal troubled particle theorists," admits Widrow. Their calculations also put the seed magnetic field at 10^{-50} gauss, meaning a powerful dynamo would be required to reach the level of magnetism we observe now.

Nevertheless, Widrow and Turner's idea proved to be a starting point for other investigators. "Their theory was the very first suggesting the production of magnetic fields during inflation," says Dominik Schleicher at the University of Göttingen in Germany. "It marked a cornerstone in our understanding."

In 2013, physicist Leonardo Campanelli of

The ASKAP radio telescope in Australia is helping to hunt for signs of cosmic magnetism

the University of Bari in Italy explained how these fluctuations may have seeded primordial magnetism without resorting to non-standard physics. He used a mathematical trick called renormalisation, something particle theorists already use to tame infinities that render their equations useless. "No one had thought to apply renormalisation to the primordial magnetism problem," says Campanelli.

His work arrives at a much larger value for the primordial magnetic field of 10^{-12} gauss. That's still smaller than the 10^{-6} gauss value measured in intergalactic space. But this background field, into which the first stars and galaxies are born, is enough to be amplified to present-day values, he says.

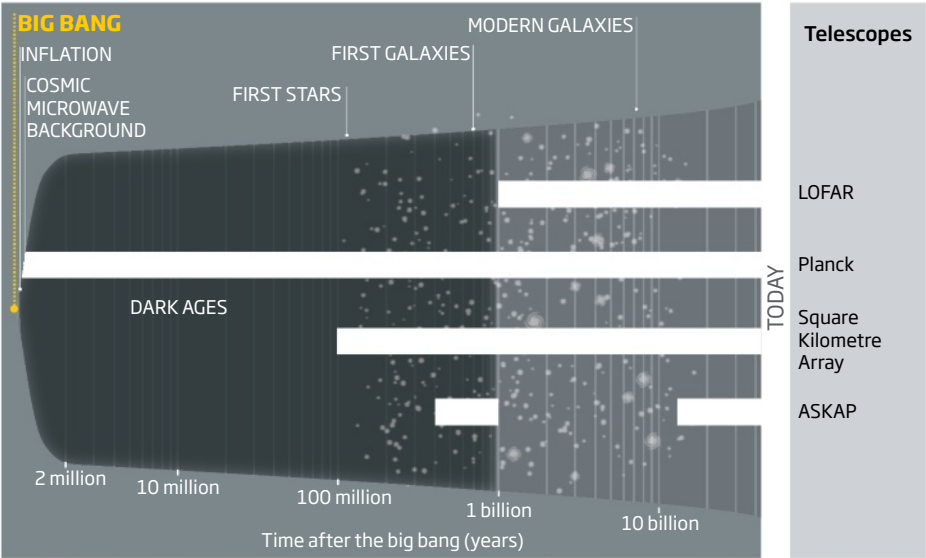
Widrow is impressed with Campanelli's paper. "If the calculations in this paper hold up, then it makes large-scale magnetic fields a natural and expected outcome of inflation, rather than something that requires exotic, if not controversial, modifications to the laws of physics," he says.

Others pinpoint problems with any theory that generates a cosmos-wide magnetic field during inflation or soon afterwards. That's because it might have been almost wiped out during an era known as the dark ages.

For the first 378,000 years of its existence

Time travellers

Telescopes are looking back in time to see magnetic fields and understand their origin and evolution



CHARLES BREWER/REX

the universe was so hot that there were no atoms, only electrons, nuclei and photons. This swirling cauldron of electrical charge was the perfect place to brew a stronger magnetic field, reinforcing the seed field present soon after inflation.

As the universe expanded, it cooled enough for protons to capture electrons and become neutral hydrogen atoms. As they combined, the particles allowed a rush of radiation to flood the universe – this is the famous cosmic microwave background (CMB).

Then came the dark ages, so called because



universe and evolved will help us to constrain models of their origins.

According to astronomer Richard Davis at the University of Manchester, UK, we could soon have the first analysis of measurements of cosmic magnetism taken by the Planck satellite, which has been studying the CMB. If primordial fields did exist when the universe was 378,000 years old, their “imprint” should be there on the CMB.

Telescopes pull together

Joining the Planck scientists are radio astronomers at the Low-Frequency Array for radio astronomy (LOFAR), whose antennas stretch over five European countries. Then there are the researchers working on two instruments in the Australian outback – the Australian Square Kilometre Array Pathfinder (ASKAP) and the Murchison Widefield Array. They are looking for radio synchrotron emissions: radio waves emitted by cosmic-ray electrons spiralling around the field lines.

LOFAR is designed especially to measure longer wavelengths, so it will be able to probe regions of weaker magnetic fields – such as between galaxies – and explore the question of how far from the galactic disc the fields extend. It will also be able to probe fields in galaxies that formed early on in the universe. Gaensler, one of the leaders of ASKAP’s project on cosmic magnetism, is confident about being able to identify which of the many theories is correct within a year or two.

If they find evidence of strong magnetic fields in proto-galaxies, this would favour the idea that the magnetism was started by shock fronts or plasma fluctuations in young galaxies, says Rainer Beck at the Max Planck Institute for Radio Astronomy in Bonn, Germany. However, if the first fields are found near galactic nuclei, this may favour early stars or early galactic dynamo action.

Greater observing power is on its way, too, in the shape of the gigantic radio telescope, the Square Kilometre Array (SKA) in Australia and South Africa. This complex of thousands of radio antennas will allow researchers to study magnetic fields at 10 times the resolution available today. The SKA is set to make its first measurements in the early 2020s. It will probe the epoch of reionisation in an effort to identify the first objects to appear in the universe. It will also be searching for early magnetic fields. “The SKA will allow us to measure the intensity and polarisation of radio waves with unprecedented sensitivity,” says Ethan Vishniac, an

POWER OF THE WEAK

It would take 10 million Milky Ways to stick a shopping list to your refrigerator door – that’s how incredibly weak our galaxy’s magnetic field is. It is still enough to influence the motion of charged particles called cosmic rays, bending their paths and even trapping them inside our galaxy for millions of years.

Without magnetic fields, cosmic rays would fly out of the Milky Way shortly after they are made, points out Harvard University astronomer Abraham Loeb. The consequences would be profound. “Cosmic rays are an important component of the Milky Way. They ionise gas deep in the proto-planetary discs. They are important for mutations in biology on Earth. In short, they are an important fact of life,” he says.

Indeed, life’s big break could be the work of high-energy cosmic rays getting their oomph from being slingshot by magnetic fields. It seems that these particles initiate the chemistry that forms sugars, amino acids and other building blocks of life found in dense gas clouds.

Despite this, we don’t know for sure where cosmic rays come from, because cosmic magnetic fields deflect them. By studying the fields, says Loeb, we will find clues about the origins of cosmic rays and solve a very important mystery.

astrophysicist at McMaster University in Hamilton, Ontario, Canada.

If the SKA finds that strong fields existed around the first objects, the primordial idea may get support, says Beck. It would indicate that magnetic fields preceded the formation of galaxies – and may have had an effect on their evolution. And in this case, help could come from Planck or the next generation of CMB satellites that researchers would like to build.

In a decade or so, once measurements from all these telescopes and satellites have been analysed, our cosmic maps will have to be redrawn. “Most of the numerical simulations of gas dynamics during the cosmological evolution of galaxies ignore magnetic fields,” says Harvard astronomer Abraham Loeb. “The next frontier will be to incorporate them, along with cosmic rays, and try to see what effects these components have on galaxies.”

When we understand how the invisible hands of gravity and magnetism operate across the cosmos, then we will know how our universe really works. ■

nothing around at the time emitted light. The only source of radiation came from hydrogen atoms, which released radio waves with a distinctive wavelength of 21 centimetres.

Perhaps problematically for cosmic magnetism, the numbers of charged particles plummeted. During the dark ages, there was just one free electron or proton for every 10,000 hydrogen atoms. Because magnetic fields rely on the movement of electrons and protons, some researchers think that the seed field might have winked out at this point.

The dark ages lasted until the first sources of

“The galaxies we observe were caused by quantum fluctuations in the moments after the big bang”

light appeared in the universe. As these stars and galaxies formed, they released huge amounts of radiation that stripped electrons from the hydrogen atoms. This epoch of reionisation lasted roughly a billion years and meant the universe was awash with magnetic-field-friendly electrons and protons.

We are not yet sure how well cosmic magnetism fared during these tumultuous times. However, after decades of exotic theories, some answers might be forthcoming.

By combining observations from several telescopes that probe the universe at different eras in cosmic history, researchers will be able to trace the evolution of magnetism. Knowing how the magnetic fields looked in the infant



Shaken *and* stirred

An unexpected cosmic ingredient is pushing mighty dark matter around, says cosmologist **Andrew Pontzen**

YOU'D think Carlos Frenk would be pleased that no one calls him a crackpot any more. He wasn't always so lucky. "I would stand up at conferences and have people almost throwing rotten tomatoes at me," he says.

His offence was to be an ardent advocate of a then controversial idea – that most of the universe's matter comes as a cold, heavy soup of invisible dark matter. Today that is the orthodoxy. Wherever dark stuff accumulates, so the standard story goes, normal matter meekly follows, irresistibly drawn in by its overbearing gravity. This matter forms stars, and then galaxies are born – meagre pricks of light in a domineering dark empire.

But the confidence of such statements now has Frenk worried. "I suddenly realised that young scientists were taking dark matter for granted, and was absolutely scandalised," he says. You can see his point. Experiments that are supposed to conjure up dark matter have so far produced nothing. Searches for its particles streaming through the Earth have

thrown up confusing, contradictory results. Models of how the stuff shapes the visible cosmos veer between triumphant confirmation and abysmal contradiction.

As a young theoretical cosmologist myself, I am among dark matter's disciples. To my mind there is just too much in the universe we can't explain without the stuff. But there is perhaps a way out of its worst dilemmas. Dark matter really does exist; we just need to rethink the idea that it holds all the power in our star-spangled cosmos.

It was about a decade ago that my undergraduate physics lecturer casually introduced me to the idea that five-sixths of the matter in the universe is invisible. Dark matter was originally invoked to explain the observation in the 1930s that clusters of galaxies whirl around too fast for the amount of ordinary matter in them. In the 1970s it was also used to explain why galaxies themselves are spinning too fast, as if subject to an extra gravitational tug. Even so, I recall thinking that you might as well base explanations

of the cosmos on magic fairy dust.

But experience made me a true believer. The way galaxies and other massive objects bend light vindicates the idea that there is more to the cosmos than meets the eye. Patterns in the cosmic microwave background, the big bang's afterglow, reveal matter in the early universe caught in a finely balanced competition between gravitational contraction and expansive pressures in a way that agrees with dark matter theory in stunning detail. In my own research on how galaxies form, to reproduce anything like the web of galaxies spun across the cosmos we need dark matter just as Frenk and others ordered it: a cold soup of stuff that barely moves at all.

Wimpy out

Pleasingly, particle physics supplies a ready recipe for this soup. The theory of supersymmetry is a favoured step beyond our current "standard model" of particles and their interactions. It holds up a mathematical mirror to reality by asserting that every particle so far discovered has a generally heavier partner. Some of these super-partners are weakly interacting massive particles, or WIMPs. These have mass (and so produce and respond to gravity) but do not interact with light (and so can't be seen). The number of WIMPs that should have been created in the big bang coincides tidily with the density of dark matter inferred from cosmological observations – a happy conjunction sometimes known as the WIMP miracle.

But do miracles really happen? No experiment that might have produced supersymmetric particles, not even the Large Hadron Collider at CERN near Geneva, Switzerland, has seen a hint of them so far. The simplest supersymmetric theories have already been ruled out, and more complex versions await their fate when the LHC restarts at a higher energy, probably in 2015. "After that, if they don't find supersymmetric particles within about a year, I think it'll be dead," says Ben Allanach, a particle theorist at the University of Cambridge. "I'll start to work on something else, and I think a lot of ➤

other people feel the same way.”

That’s not the only difficulty. Fiddly experiments looking for the fingerprint of cosmic WIMPs as they stream from space are producing highly confusing results. The DAMA experiment at the Gran Sasso National Laboratory in central Italy has seen a signal that changes on a yearly cycle. That is what we would expect if Earth is moving relative to a placid cold dark sea as it trundles round the sun – but other experiments flatly contradict the finding. Space-based missions such as the PAMELA satellite and the Alpha Magnetic Spectrometer (AMS) aboard the International Space Station have measured excesses of antimatter particles that might be produced when two WIMPs collide – but these don’t really fit our expectations. Overall, “there’s huge scepticism about the claims of dark matter detections because other experiments rule that out,” says Frenk.

Perhaps the most damaging blow, however, is that when we look at the details, WIMP-based cold dark matter doesn’t seem to be the consummate galaxy sculptor we thought. Michael Boylan-Kolchin, a cosmologist at the University of Maryland, College Park, has run

simulations of standard cold dark matter’s effect on the formation of dwarf spheroidals, mini galaxy-ettes that swarm around the Milky Way. Boylan-Kolchin could infer the dark matter content of these dwarf galaxies by watching how stars move around inside them. “It didn’t seem to make sense: things were more massive and dense in the simulation than the things we see in the real universe,” he says. Is it time to change our dark matter order?

A warmer broth

There are alternatives on the menu. If, rather than a cool gazpacho, dark matter were a hot broth of zippier particles, it would lump less readily and so form more diffuse galaxies. In the 1980s, measurements of neutrinos convinced some researchers that the collective mass of these oddball particles, which race around at close to the speed of light, would be enough to explain dark matter. But this mass turned out to be a huge overestimate, and neutrino-based hot dark matter had the opposite problem to cold dark matter: it moved around too fast to settle down into relatively compact structures like galaxies at all.

There is a third way. A while ago, Frenk set his team to work on a “Goldilocks” solution: dark matter that is not too hot, not too cold, but just right. To their surprise, they could make a lukewarm variant of dark matter produce the right, wafty dwarf spheroidal galaxies – without ruining the rest of cosmology.

That still has consequences. The leading candidate for a warm dark matter particle is a heavier, more elusive sibling of the neutrino known as a sterile neutrino. The LHC might manufacture sterile neutrinos indirectly in its particle collisions, but their signature is so subtle it wouldn’t necessarily know it had. Our best hope of spotting sterile neutrinos is when normal neutrinos spontaneously turn into them and disappear off a detector’s radar.

So if warm dark matter is the solution, experiments such as DAMA have been looking in the wrong place. “From an experimental point of view it would be very tragic because there’s been a huge investment, particularly in the direct detection of WIMPs,” says Frenk. The idea has consequences for theorists, too: warm dark matter is divorced from the sort of particles predicted by supersymmetry, leaving both ideas considerably weaker.

In the meantime, Jorge Peñarrubia of the University of Edinburgh, UK, and his colleagues have been looking at how dark matter is distributed within nearby dwarf spheroidal galaxies. It seems to be evenly spread across their diameter. “This constant density was something we did not expect,” Peñarrubia says. Simulations with any temperature of dark matter – cold, hot or warm – produce dwarf galaxies that are more densely packed towards their centres. This discrepancy between theory and observation repeats itself in slightly larger, more distant galaxies.

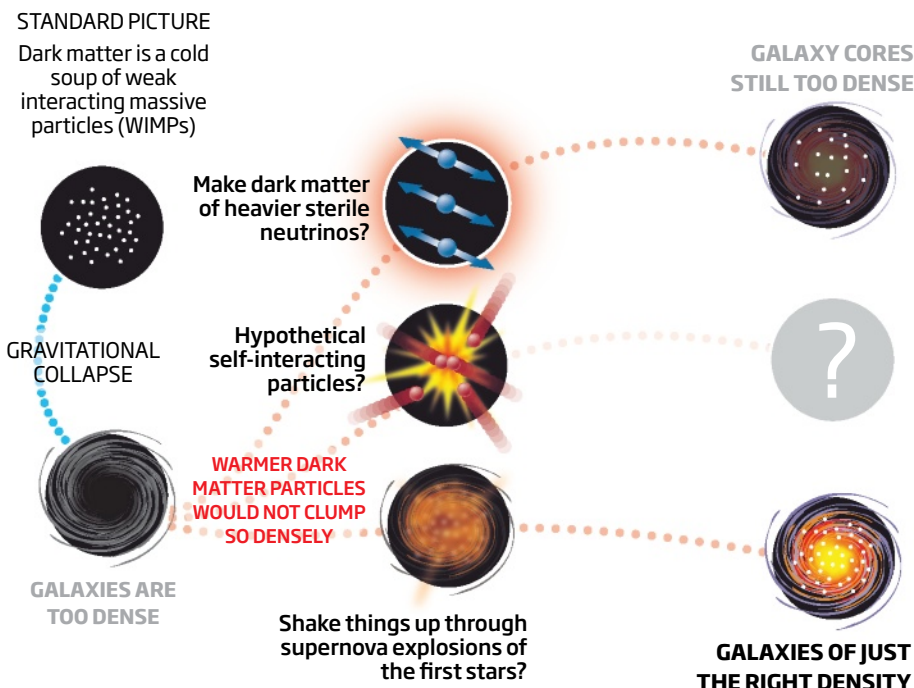
Some exotic flavours of dark matter might help out, such as “self-interacting” dark matter, feisty particles that constantly ricochet around and so resist being sculpted into tight central cusps. But we are at a loss as to what hypothetical particle would bounce around just the right amount in dwarf galaxies and not too much in galaxy clusters. To me, this particular soup smells a little fishy.

So we are at an impasse. Cold dark matter does not quite do all the jobs we ask of it – but then again, nor does anything else.

My own hunch is that, oddly, cold dark matter might be the right stuff after all. The price we must pay is to stop assuming that it is the totalitarian force in the governance of galaxies. Stars generate huge amounts of energy in their lifetimes. When their time is

Sculpting galaxies

The gravitational collapse of dark matter is thought to be the seed of galaxy formation. But to make dwarf galaxies that look true to life, we need dark matter to be a warmer soup



up, they explode in supernovae. Gas spiralling into black holes generates vast amounts of heat. The energy from either of these sources is enough to send enormous quantities of gas swirling violently around inside a galaxy. Dark matter is not immune to these huge gravitational ructions: it begins to move in concert. Simulations I and a number of colleagues have been performing over the past few years suggest that, if the normal gas is shaken enough, it sends dark matter into a real funk, swirling it around like snowflakes in a snow globe.

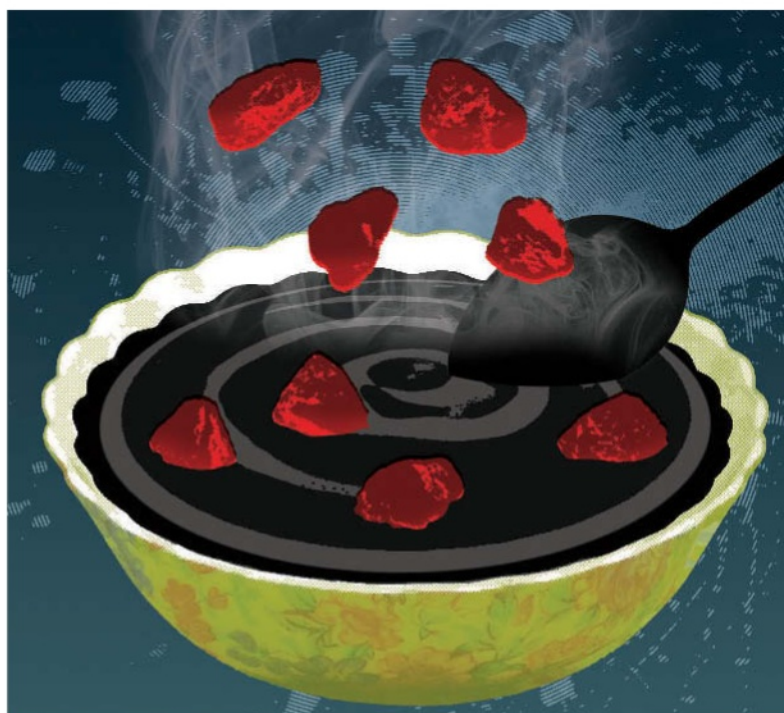
Dark matter particles could then be cold and supersymmetric again, and simply get hot under the collar when bullied by exuberant normal matter. The increased energy of this protogalactic soup stops it coalescing too densely, so the structure of the Milky Way's dwarf satellites makes sense again. The only remaining puzzle is why direct searches for dark matter have produced such ambiguous results so far. Perhaps cutting-edge particle physics experiments are just hard.

Peñarrubia thinks it a credible enough picture. "The redistribution of dark matter requires vast amounts of energy, and supernova explosions are the only plausible source," he says. But he cautions against thinking that supernovae are enough on their own, especially to produce dwarf spheroidal galaxies of the right density. "The small number of stars in these galaxies limits the supernova energy to values that are only just sufficient," he says. Boylan-Kolchin is also sceptical that jumbling everything up with cosmic explosions solves all the problems. "My personal feeling is that it's unlikely that these effects can save cold dark matter theory," he says.

Hidden signals

Of course I think he's being pessimistic – and would point to other, albeit highly tentative, evidence for the idea of rambunctious normal matter kicking dark matter around. It comes from NASA's gamma-ray satellite telescope Fermi, which has been on the lookout for visible signals given off by dark matter since 2008. That might seem futile; the whole point is that you can't see dark matter. But even cold dark matter theory suggests you can see indirect signs of the stuff if, for instance, its particles collide and annihilate in a puff of energy and a very visible flash of gamma-ray light.

The Milky Way's dwarf satellites are again a good place to look. They are close by, and naturally very dim, meaning any gamma-ray



"If the gas shakes enough, it sends dark matter into a real funk, swirling it around like snowflakes in a snow globe"

photons from them are likely to come from dark-matter annihilations. Officially, we haven't seen any yet. "The bottom line is that so far we have not detected any signal," says former Fermi team member Andrea Albert, now at the SLAC National Accelerator Laboratory in California.

In April 2012, however, a researcher from outside the Fermi collaboration, Christoph Weniger of the Max Planck Institute for Physics in Munich, Germany, set astrophysics afizz with the suggestion that Fermi had indeed seen a dark matter signal – not from dwarf galaxies, but from near the heart of the Milky Way itself. His work indicated that a significant excess of gamma rays from that general direction, all with the same energy of around 130 gigaelectronvolts, was buried in publicly released data from the Fermi satellite.

The signal is weak, and where it is strongest is not at the Milky Way's centre, where you would expect dark matter to accumulate, but around 1 degree off-centre. This misalignment nourishes a suspicion that it is the product of some obscure miscalibration of the telescope.

Alternatively, normal matter might be

buffeting the dark stuff around. That at least is the implication of simulations by Mike Kuhlen from the University of California, Berkeley. "It surprised me, but you almost immediately see the dark matter gets offset in the simulation when gas and stars are included," he says.

So, all is well with standard cold dark matter, as long as you factor in the effects of normal matter. Not so fast, says Frenk. If supersymmetric particles annihilating each other were the source of the gamma rays, they would be producing them not with one standard energy, but with a spread: the annihilation mechanism generates electrons and positrons, which gradually give up their energy in unpredictable fits and bursts. "The case is absolutely fascinating, but I don't think we've found anything yet," he says.

Things might change in a moment if the many experiments looking for dark matter were to start producing consistent results. But that will take years at best. In the meantime, the discord is music to Frenk's ears. "We don't know whether cold dark matter's right," he says. "If everyone just buys into an idea, things don't progress." ■



Black hole universe

Legions of black holes drifting around the cosmos may at last explain dark matter, says **Marcus Chown**

MOST of the matter in the universe gives out no light, or at least so little that it is currently undetectable. Yet we know it is out there because its gravity keeps stars and galaxies in their orbits. Pretty much everyone thinks that this so-called dark matter is made of hitherto undiscovered subatomic particles. Physicists are hopeful they will find a candidate in high-speed collisions at the Large Hadron Collider at CERN, near Geneva, Switzerland.

But could we have got dark matter all wrong? Mike Hawkins thinks so. He believes that rather than particles, what we call dark matter is actually legions of black holes created shortly after the big bang.

It is a controversial claim. Yet Hawkins, who is an astronomer at the Royal Observatory in Edinburgh, UK, believes he has persuasive evidence gathered over years of observations. If he is right, it radically changes our picture of the cosmos. "We live in a universe dominated by black holes," he says.

The story dates back to 1975, when Hawkins

began monitoring a portion of the southern sky, night after night, using the UK Schmidt Telescope at Siding Springs in Australia. After five years, in addition to the stars he was looking for, he found something unexpected: thousands of objects that were brightening and dimming extremely slowly. They turned out to be quasars, the super-bright cores of newborn galaxies at the edge of the universe.

Yet Hawkins was puzzled. Quasars are powered by matter that shines fiercely as it swirls into a supermassive black hole typically billions of times the mass of the sun. Their light does vary, for example as stars are ripped apart by the black hole's enormous gravity. But these flare-ups make quasars brighten and fade within a few days, not years and decades as Hawkins found.

He wondered if the variability was down to a phenomenon called microlensing. If a massive body happens to pass between us and



“What we call dark matter may actually be legions of black holes created shortly after the big bang”

a quasar, its gravity will bend the quasar's light and magnify it. Perhaps that was responsible for the light variation, rather than it being something intrinsic to the quasar itself. Was it sufficient to explain the decade-long variation Hawkins was seeing?

His calculations showed that it could, provided that the bodies passing in front of the quasar had a mass roughly that of the sun. So what could these massive, microlensing bodies be?

The obvious answer is stars. However, they are ruled out for a compelling reason that dates back to the first few minutes after the big bang. The universe is the way it is now because ever since then it has consisted of a certain ratio of photons to protons and neutrons, which are collectively called baryons.

We can tell what this ratio is from the cosmic microwave background – radiation left over from the big bang. It turns out that there can be at most only twice as many baryons in the universe as we see tied up in stars and galaxies. That's a problem for microlensing of quasar light, because to see the effect with every quasar, which we do, means the bodies bending the light around must add up to far more than double the number of baryons we see in the universe. “There just isn't enough baryonic matter for the lenses to be ordinary stars,” says Hawkins. “They must be dark matter.”

The baryonic argument rules out many other possible candidates that belong to a family called MACHOs, or massive astrophysical compact halo objects. MACHO is a catch-all name for any dark body that might account

for dark matter, and includes neutron stars, black holes and stars so small that they give out little or no light.

But if everything made of ordinary matter is ruled out, that leaves Hawkins with only one candidate: small black holes weighing about the mass of the sun and not much bigger than a village. Stephen Hawking at the University of Cambridge and others have said that these could have formed spontaneously when the universe was about 1/100,000th of a second old.

Star quality

At this time, the building blocks of matter known as quarks went from being free to condensing, like droplets forming in steam. The “droplets” were protons, neutrons and a plethora of other subatomic particles that can only be created in collisions between high-energy particles today.

This process would have resulted in isolated super-dense regions of particles that could have shrunk under their gravity to form primordial black holes. “The characteristic mass of such holes is about the mass of a star,” says Hawkins, “exactly what is required to be quasar microlensing candidates.”

Hawkins has had a tough time convincing other astronomers. Although no one knows what dark matter is made of, most researchers now favour particles called neutralinos, which are quite unlike ordinary baryonic matter. One

reason is that our most promising theories of particle physics points to neutralinos as dark matter. Another is that a wide-ranging search for MACHOs has largely drawn a blank.

Over a period of seven years, an international team called the Macho Collaboration monitored 3 million stars in the Milky Way's satellite galaxies, the Large and Small Magellanic Clouds, to see if there were any signs of MACHOs. The logic was simple. The dark matter in our galaxy is believed to be in the form of a vast spherical halo. If the halo is made of MACHOs, then occasionally one will pass in front of a star in one of the satellite galaxies and microlense its light.

The results were not promising for black holes. The team saw just 17 microlensing events and concluded that MACHOs – whatever they might be – make up less than 20 per cent of the mass of the halo. Later surveys have reported an even smaller percentage. Hawkins, though, believes the results were mistakenly interpreted, and that the microlensing events are consistent with a halo made entirely of MACHOs. “There are many uncertainties with these projects, including the structure of the halo and detection efficiency,” he says.

He points to several strands of evidence in favour of primordial black holes. For a start, everyone agrees that quasar light definitely varies over periods of many years. Sometimes, the light from a quasar is lensed by an intervening massive galaxy, and this can bend the light along several paths to form multiple images of the quasar. If the light variation is intrinsic to the quasar, any change in one image should be matched by an identical change in the rest.

Although this is occasionally seen, the images more often vary independently. “This can only be explained by microlensing,” says Hawkins, describing how a primordial black hole can drift into the path of light that goes on to form one image but not the others (see diagram, page 64).

A second strand of evidence comes from the timescale with which the light brightens and dims. If this is a feature of the quasar itself, then quasars farther away should vary over a longer period than closer ones. This is ➤

because things happen more slowly to distant objects, an effect called time dilation.

But according to Hawkins's analysis, there is no difference in variability between distant and nearby quasars. "This is compatible with the variation being caused by microlensing," he says, "and not compatible with any form of intrinsic variation."

Bernard Carr of Queen Mary University of London shares Hawkins's enthusiasm for primordial black holes.

"I think his results deserve serious consideration," he says.

Still, no matter how much evidence Hawkins puts forward, he faces a problem. Quasars are complex beasts and our lack of understanding of them leaves the door wide open for alternative theories. For example, some have suggested that their variability

"If black holes do account for dark matter, they might have survived a series of big bangs and big crunches"

could be due to changes in the rate at which matter flows into the black hole. "It's difficult to know how to counter such an unfalsifiable argument," Hawkins says.

Hawkins's hypothesis, on the other hand, is falsifiable. If one of the many experiments searching for particles of cosmic dark matter finds some, then it will be curtains for primordial black holes. So far there are some intriguing claims. The oldest comes from researchers working on the DAMA experiment, at the Gran Sasso underground laboratory in Italy. They claimed in 2000 to have found evidence for dark matter particles raining down on Earth. An improved version of their experiment gave results that agreed with the evidence. A second experiment called CoGeNT, deep in the Soudan mine in Minnesota, also claims to have seen the effect.

Both claims are controversial, however, and several other experiments have failed to see anything. "If this holds up, then dark matter is indeed made of particles and I am wrong," Hawkins acknowledges.

What he really needs is a killer observation, but that won't be easy. If the halo is made of primordial black holes, there will trillions of them, greatly outnumbering stars in the galaxy. But they will be at least 40 light years apart – 10 times the distance between the sun and its nearest neighbour, Alpha Centauri. Add to that their smallness – less than about 3 kilometres across – and finding a halo of black holes is worse than looking for a needle in the proverbial haystack.

The best way to find black holes, if indirectly, is in another microlensing survey. "The tragedy of the Macho Collaboration's experiment is that it stopped too soon," says Hawkins. "We really need to start the search again."

If he is right, and stellar-mass black holes do account for the universe's dark matter, they need not necessarily have been created in the very early universe. Carr and his colleague Alan Coley at Dalhousie University in Halifax, Nova Scotia, Canada, have been examining models which have the universe going through a series of big bangs and big crunches. In these, black holes created at an earlier time might survive these repeated bounces and still be with us today. "Macroscopic black holes surviving from the last cycle of the universe would be observationally indistinguishable from ones forming in this cycle," says Carr. "So they might indeed induce quasar lensing."

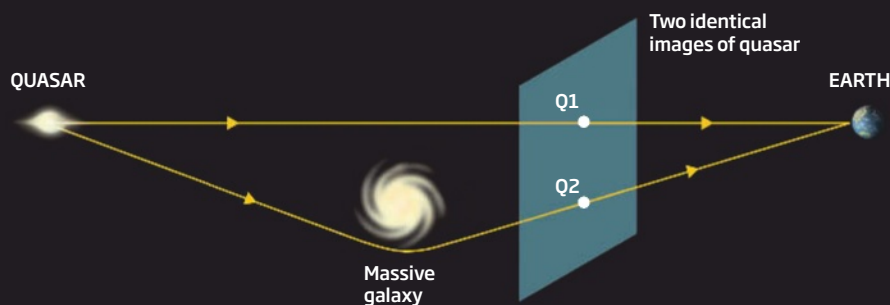
We have long suspected that dark matter is quite unlike anything we come across here on Earth. Perhaps much of it wasn't even made in our universe. ■

Lens on mini black holes

Light from a distant quasar is bent by the gravity of a massive galaxy to produce two bright images. How the images look depends on other objects present

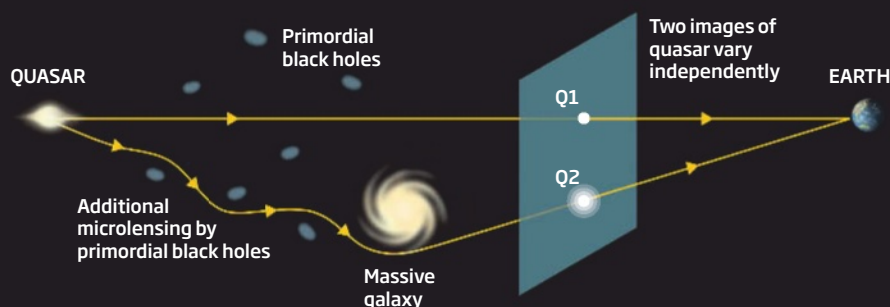
GRAVITATIONAL LENSING

Any variation in the quasar light shows up the same in both images



GRAVITATIONAL LENSING PLUS MICROLENSING

Primordial black holes bend and magnify light following a particular path





Royal Society prize lectures 2014

**Join us to celebrate outstanding science
delivered by award winning scientists**

Croonian Lecture

12 June 6.30pm

Given by Professor Brigid Hogan FRS

GlaxoSmithKline Prize and Lecture

23 September 6.30pm

Given by Dr Nicholas Lydon FRS

Clifford Paterson Lecture

27 October 6.30pm

Given by Professor Polina Bayvel FREng

Francis Crick Lecture

4 November 6.30pm

Given by Dr Duncan Odom

The Rosalind Franklin Award Lecture and the Milner Award Lecture will also take place in 2014. Details will be published in our events email bulletin and on royalsociety.org/events

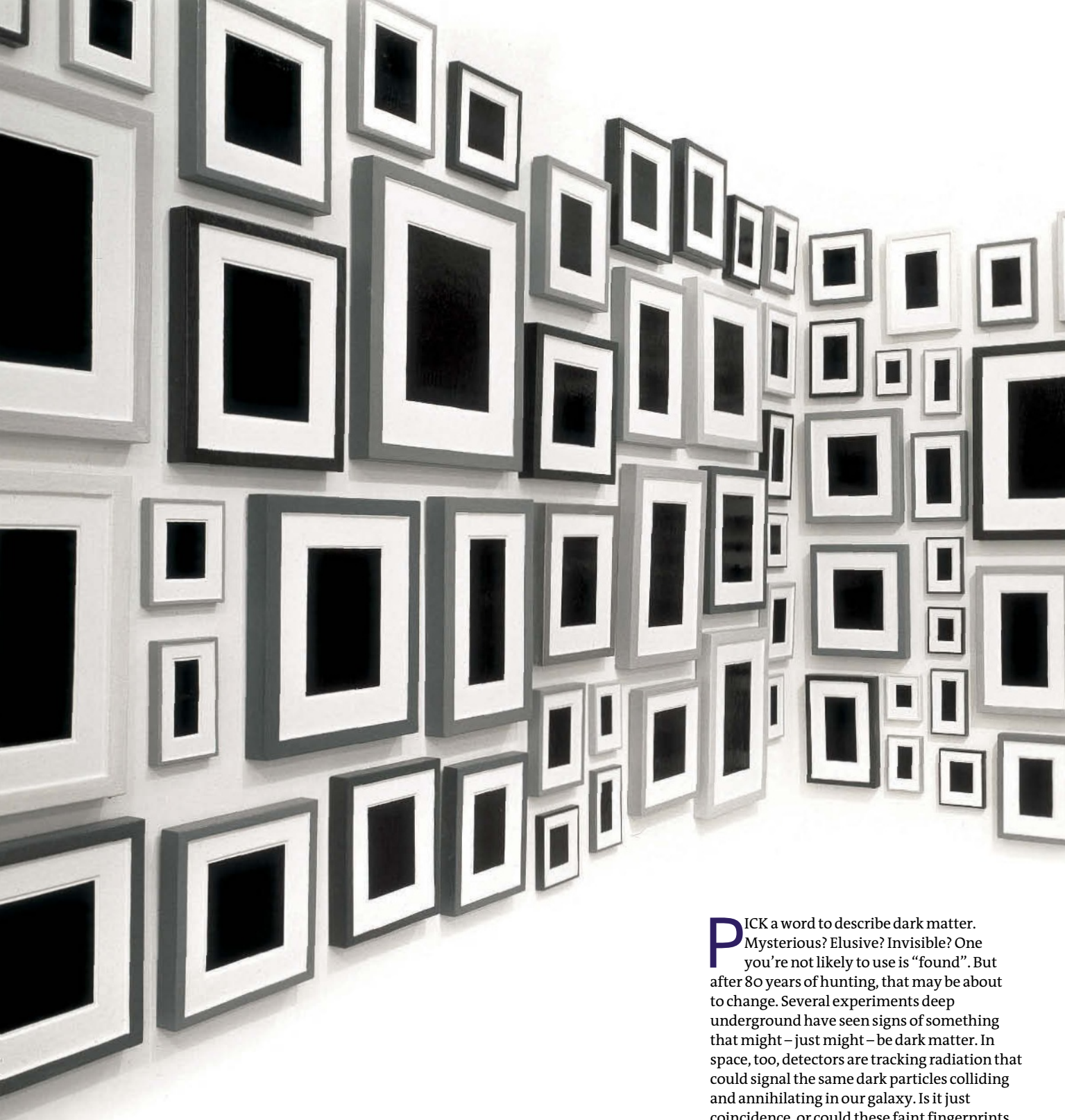
The Royal Society, 6 – 9 Carlton House Terrace, London, SW1Y 5AG

Free admission – doors open at 6pm

For more information visit: royalsociety.org/events

THE
ROYAL
SOCIETY

Registered Charity No 207043



PICK a word to describe dark matter. Mysterious? Elusive? Invisible? One you're not likely to use is "found". But after 80 years of hunting, that may be about to change. Several experiments deep underground have seen signs of something that might – just might – be dark matter. In space, too, detectors are tracking radiation that could signal the same dark particles colliding and annihilating in our galaxy. Is it just coincidence, or could these faint fingerprints really all be from the same dark hand?

Dan Hooper, a theoretical astrophysicist at the University of Chicago, believes that we may already have glimpsed dark matter. "I happen to be in the relatively rare minority of my colleagues who think we probably have," he says. "I'm not certain – I just think it's likely." If Hooper's hunch is right, the findings challenge our understanding of what dark matter is. The stuff that makes

ALLAN MCCOLLUM: PLASTER SURROGATES, 1982/84. ENAMEL ON CAST HYDROSTONE, SIZES VARIABLE, EACH UNIQUE. INSTALLATION: RICHARD KUHLENSCHMIDT GALLERY, 1984 (DETAIL)



Portraits of darkness

Are we finally on the verge
of unveiling dark matter?
Robert Adler investigates

up 85 per cent of matter in the universe might be even stranger than we imagined.

Dark matter earned its name because it gives out no light for our telescopes to catch. Yet we see its gravitational footprints everywhere we look, from ripples in the cosmic microwave background to the largest galaxy clusters. If the big bang hadn't created dark matter in abundance, there wouldn't be enough gravity for gas and dust to coalesce into stars and galaxies. Remove dark matter and the spinning Milky Way would fly apart. "Dark matter is the scaffolding upon which galaxies are built," says James Bullock, a cosmologist at the University of California, Irvine. "Without it, we really don't think we would have much structure in the universe."

Astronomers discovered it, cosmologists map it, but it is particle physicists who most keenly want to understand it. They know it is something fundamentally different,

beckoning from outside the standard model that describes the particles and forces of the visible universe. Many researchers believe that dark matter is the window to an uncharted world. But what is it?

The leading candidate for dark matter is a WIMP, or weakly interacting massive particle. WIMPs get their name because we suspect that they interact with visible matter only through gravity and the weak nuclear force. Slight as the weak interaction is, it should still make WIMPs detectable. In the decades-long search for them, dark matter hunters have taken three complementary approaches: direct detection, indirect detection and colliders.

The Cryogenic Dark Matter Search (CDMS) typifies direct detection – and it is one of the experiments that may have found dark matter. The idea is that every so often one of the billions of WIMPs whizzing through the galaxy interacts with the nucleus of an atom

in the detector, giving it a tiny kick. The CDMS sandwiches targets made from silicon and germanium between sensors that register electrons and the faint vibrations produced in those collisions. To protect their detector from background radiation that might mimic a signal, the CDMS team has buried it 700 metres underground in a mine in Minnesota. They also chill it to one-hundredth of a degree above absolute zero to minimise another copycat signal from thermal noise.

In April 2013, the CDMS team announced three collisions in silicon as likely WIMPs. These weighed in at 8 gigaelectronvolts (GeV), or around eight times the mass of a proton. But with just three collisions, the researchers can't be sure.

Still, the CDMS findings added an extra frisson to the dark matter hunt. That's because they are in line with an experiment called CoGeNT, also located in the Minnesota mine, ➤



"The discovery of dark matter would be the find of the century. So why aren't physicists celebrating?"

that had seen particles with the same mass in 2011. Like CDMS, CoGeNT didn't see them in sufficient numbers to claim an outright discovery, yet taken together the findings are tantalising. Especially when you consider that one of CoGeNT's original goals was to settle an issue that is infamous in dark matter circles, known as the DAMA/LIBRA anomaly.

The DAMA/LIBRA experiment lies 1400 metres beneath Italy's Gran Sasso National Park and has been running for more than a decade. Its detectors measure a striking signal that ebbs and flows with the seasons, exactly as you would expect if our planet were sailing through a sea of dark matter particles on its way around the sun. DAMA/LIBRA implicates particles with a mass of around 10 GeV, similar to the mass of those found by CDMS and CoGeNT (see diagram, below). Unlike its counterparts, there's no doubting the strength of the DAMA/LIBRA signal – in physics parlance, it has a standard deviation of 8.9 sigma, placing it well above the 5-sigma level considered as a discovery.

Yet for many years, no one believed the DAMA/LIBRA result was down to dark matter, because the experiment was one of a kind. Then came CoGeNT. Juan Collar at the University of Chicago, who leads the CoGeNT team, was as surprised as anyone when it didn't rule out the DAMA/LIBRA findings. "Rather than excluding the DAMA halo, we found an excess at low energies that we can't explain away, and hints of an annual modulation," he says.

The discovery of dark matter would be the find of the century. So why aren't physicists celebrating? One catch is that DAMA/LIBRA records far more strikes than other experiments expect. That's a problem that needs to be solved before other researchers will accept the team's claims. For James Boyce, an experimental physicist at Jefferson Lab (JLab) in Newport News, Virginia, the need for caution matches the significance of the goal. "Claiming dark matter has been found is such a serious and monumental step in our exploration of the universe, we dare not trivialise the discovery," he says.

Another issue is that for decades theorists have favoured heavier WIMPs, in the 100 GeV range. Such heavy WIMPs fall out naturally

from supersymmetry and other theories that attempt to go beyond the standard model. A lightweight WIMP ran counter to many of their expectations.

However, the biggest problem is that the world's largest and most sensitive direct detection experiment, XENON, has seen nothing, casting doubt on all these findings. Like an overzealous cleaning crew, the XENON team has swept away one potential observation after another, including most 10 GeV WIMPs. As they say in the polite world of physics, there's "tension" between XENON's negative results and those other experiments.

Exclusion zone

Elena Aprile, who leads the XENON collaboration, feels that tension acutely. "It's a nightmare to exclude and exclude and never find something," she says. She hopes her team's results will challenge other experimenters to build bigger and better detectors to find or

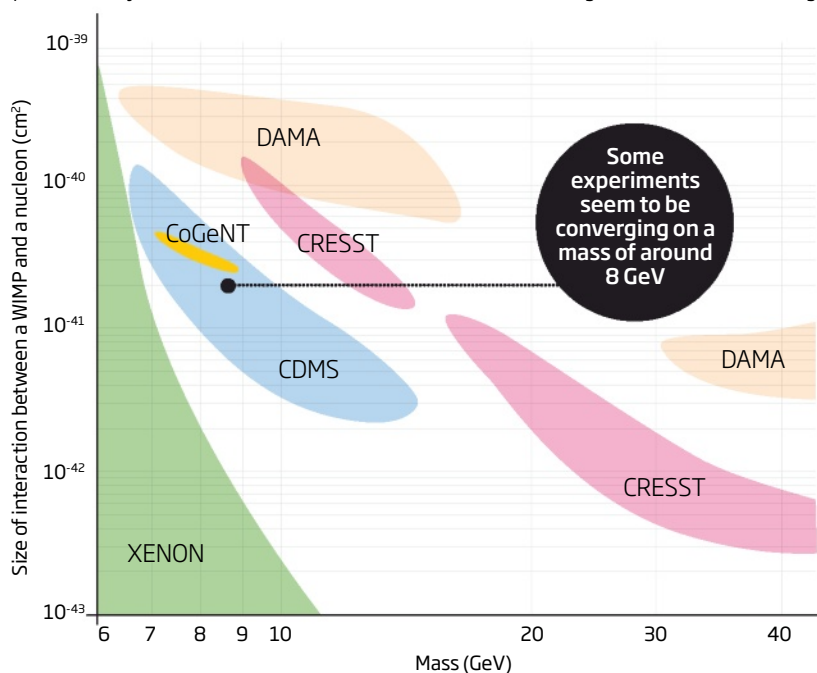
rule out particles with lower mass or even weaker interactions. Her own team is building a version that will be 100 times more sensitive than the existing XENON detector. It is scheduled to start in 2015.

Still, researchers who think a WIMP lurks in the 7 to 10 GeV range haven't given up. They know that underlying every experiment's conclusions are theoretical assumptions about how WIMPs interact with different nuclei, about dark matter's distribution and velocity in our galaxy, and just how each detector functions. Hooper has argued, for example, that if XENON's detector responded slightly differently to what was expected, its results would actually support the other experiments. He hopes that a huge project called LUX, which took the title of world's most sensitive dark matter detector when it started running in 2013, will resolve the controversy. In February 2014, results from LUX ruled out the existence of low-mass WIMPs.

However, the lightweight WIMP scenario has been boosted by events happening out in space. Several instruments are hunting for signs of dark matter particles smacking into each other and annihilating. These indirect searches look for a distinctive pattern of gamma rays or antimatter forged from the annihilation energy. One of the most promising sightings comes from the Fermi Gamma-ray Space Telescope. In 2010, it found two huge gamma-ray-producing bubbles that tower 25,000 light years above and below the Milky Way's disc of stars. Although other sources are possible, Hooper and Tracy Slatyer,

Diary of a WIMPy particle

Several dark matter experiments are looking for signs of particles called WIMPs and narrowing down their possible properties. Theory tells us what combination of mass and interaction strength can account for these signs



a theoretical physicist at the Massachusetts Institute of Technology, think the bubbles show dark particles in the 7 to 10 GeV range annihilating. “This 10 GeV-ish window is starting to be pretty suggestive,” Hooper says.

But indirect searches have their fair share of disappointments, too. Another Fermi signal once hailed as a “smoking gun” faded when more measurements were made. Early signs from the PAMELA satellite and the Alpha Magnetic Spectrometer (AMS) on the International Space Station suggested that dark matter particles are heavyweight brutes, while results from another gamma ray satellite are suggestive of WIMPs a thousand times lighter.

Ultimately, if WIMPs do exist, it should be possible to make them by smashing together particles at high energies. That means they should turn up at colliders like the Large Hadron Collider near Geneva in Switzerland. So far, however, they have been a no-show.

With the picture so bewildering, everyone agrees that what they need is more time to collect data and to build bigger and better detectors. “We want to be 100 per cent sure about what we’re saying,” says Roberto Battiston, deputy spokesperson for the AMS.

While experimentalists like Aprile describe failing to find WIMPs as a “nightmare scenario”, many theorists find the field’s confusion liberating. “To solve the dark matter mystery, we have to think as broadly as possible about what dark matter could be, so we don’t miss what it is,” says Kathryn Zurek, a theorist at the University of Michigan in Ann Arbor.

Zurek and others are exploring a much richer dark sector, with multiple dark particles and interactions and even dark atoms, dark chemistry, and ultimately a mirror universe matching ours particle for particle and force for force.

The simplest models incorporate a new dark force that lets dark particles interact, meaning that the dark sector would include dark electromagnetism and dark light, an idea first explored by Celine Boehm at Durham University, UK. In parallel with quantum theory, which calls for photons mediating the electromagnetic force, Boehm’s idea implies dark photons. “The dark photon would be a force of nature beyond the ones we know about right now,” says Rouven Essig, a particle physicist at the State University of New York in Stony Brook.

A dark force has major implications, from how visible and dark matter emerged from the big bang to how galaxies and larger structures formed. In a 2008 paper, Slatyer, along with theorists Nima Arkani-Hamed, Douglas Finkbeiner and Neal Weiner, argued that a long-range dark force could boost the rate of dark matter collisions and neatly account for DAMA/LIBRA, PAMELA and related findings.



“The dark sector could be much richer, with dark atoms, dark forces and even a mirror universe matching ours”

“It’s totally reasonable to assume that the dark sector has its own set of forces,” says Weiner. “You put that in and you find it does many things simultaneously, which is a bonus.”

That opens the door to a dazzling array of possibilities. David Kaplan, a theorist at Johns Hopkins University in Baltimore, Maryland, notes that a dark force implies dark charges, which could produce dark atoms, hydrogen-like chemistry, and perhaps even dark stars. In 2013, Andrey Katz and his colleagues at Harvard University even suggested that some dark atoms would cool and settle into galactic discs. That “double-disc dark matter” could explain several experimental anomalies and should be easily testable.

Dark light

Dark antimatter is a possibility too. According to a theory put forward by Kaplan, Zurek and Markus Luty, at the University of California, Davis, a link between normal antimatter and dark antimatter in the early universe could neatly explain the amount of dark matter we observe now.

Perhaps most exciting of all, it might be possible to make dark light in the lab. Theorist Bob Holdom at the University of Toronto in Canada showed in 1986 that dark photons can interact with normal photons and electrons, leading some researchers to think it might be easier to find dark photons than heavier dark particles. “The dark photon can be quite light and can be seen by existing accelerators,” says experimentalist Harald Merkel at Johannes Gutenberg University in Mainz, Germany.

That prospect has inspired a new raft of experiments across Europe and the US. The Mainz Microtron experiment, which Merkel works on, has started searching for dark

photons but has not yet seen one. “If we had found something, you would know,” says Merkel. “It would be as important as the discovery of the Higgs boson. It would be the beginning of the new physics.”

On the other side of the Atlantic, three experiments – DarkLight, APEX and the Heavy Photon Search – have passed preliminary tests at JLab. Each one is sensitive to a different mass range, so the three will complement each other. The Heavy Photon Search will be the first to start taking data in 2015, when JLab’s accelerator fires up after an upgrade, followed by APEX and DarkLight. John Jaros, a high-energy particle physicist at JLab, cautions that even if these experiments discover a dark photon, finding how it links to dark matter will still be challenging.

Such frenzied activities beg a question: what will it take to declare dark matter found? Researchers agree that a convincing discovery of a dark matter particle or force will require at least two different experiments whose findings match perfectly in terms of mass and willingness to interact with normal matter, and which have high statistical reliability. Collar adds two caveats. “Ideally, compatible ‘man-made’ dark matter should be found at accelerators,” he says. “And the cherry on top would be a theory that generates the particle in question, and yields other testable predictions.”

Whatever the outcome of these experiments, dark matter is spurring what physicist Rafael Lang at Purdue University in West Lafayette, Indiana, calls a new Copernican revolution. If a dark matter particle is found, it will be the first piton anchoring a long, hard climb into the dark sector. If it’s not found – and soon – everything is up for grabs, from how gravity works to the whole of physics beyond the standard model. ■

Chasing shadows

Whatever dark power is pushing the universe apart, we are hot on its trail, says Stephen Battersby





IT IS 15 head-scratching years since we noticed that some mysterious agent is pushing the universe apart. We still don't know what it is. It is everywhere and we can't see it. It makes up more than two-thirds of the universe, but we have no idea where it comes from or what it is made of. "Nature has not been ready to give us any clues yet," says Sean Carroll, a theoretical physicist at the California Institute of Technology in Pasadena.

We do at least have a name for this most enigmatic of beasts: dark energy. Now the hunt for it is really on. Astronomers have now begun a new sky survey to look for signs of the stuff among exploding stars and ancient galaxy clusters. A pack of space missions and gigantic Earth-based telescopes will soon join the chase. Meanwhile, some physicists are pursuing an unorthodox idea: that we might snare dark energy in the lab.

As yet, our knowledge of the quarry is desperately scarce. It is limited to perhaps three things. First, dark energy pushes. We first noted that in 1998, in the unexpected dimness of certain supernova explosions which told us they were further away than we expected. Space seems at some point to have begun expanding faster, as if driven outwards by a repulsive force acting against the attractive gravity of matter.

Second, there is a lot of the stuff. The motion and clustering of galaxies tells us how much matter is abroad in the universe, while the cosmic microwave background radiation emitted 380,000 years after the big bang allows us to work out the total density of matter plus energy. This second number is much bigger. According to the latest data, including microwave observations from the European Space Agency's Planck satellite, about 68 per cent of the universe is in some non-material, energetic, pushy form. That works out at about 1 joule per cubic kilometre of space.

Third, dark energy makes excellent fuel for the creative minds of physicists. They see it in hundreds of different and fantastical forms.

The tamest of these is the cosmological constant, and even that is a wild thing. It is an energy density inherent to space, which within Einstein's general theory of relativity creates a repulsive gravity. As space expands

there is more and more of the stuff, making its repulsion stronger relative to the fading gravity of the universe's increasingly scattered matter. Particle physics even seems to provide an origin for it, in virtual particles that appear and disappear in the bubbling, uncertain quantum vacuum. The trouble is that these particles have far too much energy – in the simplest calculation, about 10^{120} joules per cubic kilometre.

This catastrophic discrepancy leaves room for a menagerie of alternative theories. Dark energy could be quintessence, a hypothetical energy field that permeates space, changing over time and perhaps even clumping in different places. Or it might be a modified form of gravity that repels at long range, or an illusion born of Earth's position in the cosmos. Maybe dark energy could take the form of radio waves trillions of times larger than the observable universe – or something even more exotic (see "Arcane energies", page 72).

"Many clever people have tried to devise something better than the cosmological constant, or understand why the cosmological constant has the value it does," says Carroll. "Roughly speaking, they have failed."

The dark is rising

One way to cut to the chase might be to find out whether dark energy is changing over time. If it is not, that would exclude the cosmological constant: as an inherent property of space, its density should remain unchanged. In most models of quintessence, by contrast, the energy becomes slowly diluted as space stretches – although in some it actually intensifies, pumped up by the universe's expansion. In most modified theories of gravity, dark energy's density is also variable. It can even go up for a while and then down, or vice versa.

The fate of the universe hangs in this balance (see page 126). If dark energy remains steady, most of the cosmos will accelerate off into the distance, leaving us in a small island universe forever cut off from the rest of the cosmos. If it intensifies, it might eventually shred all matter in a "big rip", or even make the fabric of space unstable here and now. ➤

ARCANE ENERGIES

Energy is ignorance, say Salvatore Capozziello and Orlando Luongo of the University of Naples, Italy - and that could explain where dark energy comes from (see main story). If different eras of the universe are in entangled quantum states, when we try to measure some property of the early universe such as its expansion speed, quantum theory puts a limit on the information we can extract. Information loss is intimately linked to a rise in entropy, which implies energy sloshing around the cosmos. "It is possible to interpret this as dark energy," says Capozziello. By his calculations, our cosmic energy of ignorance would have the right properties to cause the real acceleration of the universe.

Or perhaps dark energy is a hologram. The holographic principle, devised to help link gravity with quantum mechanics, says that nothing can contain more energy or information than a black hole. A black hole's total energy increases with its circumference, not its volume, so perhaps the universe's energy is related to the length of some boundary akin to a black hole's event horizon. "It is an appealing idea because you get the right energy scale," says one of the idea's originators, Stephen Hsu at Michigan State University. This could reduce the excessive energy of quantum particles popping out of the vacuum to the level of dark energy. Even so, there are all sorts of technical issues, he says - not least that it is difficult to define such a boundary of the universe.

Meanwhile, Edmund Copeland and colleagues at the University of Nottingham, UK, think an unimaginable tumult of dark energy could sit beneath our serene existence. In an obscure paper published in 1974, they found the most general mathematical form of scalar-tensor theories, in which an added energy field allows the strength of gravity to vary from place to place. When they adapted these equations to produce an accelerating universe like ours, they found it includes a huge amount of vacuum energy - perhaps more than 10^{60} joules of energy per cubic metre of space. We don't feel the powerful anti-gravity effect of this dark energy because it is almost entirely blocked by the added field. If that is right, you can hold enough dark energy between your hands to disintegrate a million galaxies.

Our best estimate today, based mainly on supernova observations, is that dark energy's density is fairly stable. There is a suggestion that it is increasing slightly, but the uncertainties are too large for us to worry about this increase just yet.

The Dark Energy Survey, an international project that started collecting data in September 2013, aims to tighten things up. It uses the 4-metre-wide Víctor M. Blanco telescope at the Cerro Tololo Inter-American Observatory in Chile, attached to a specially designed infrared-sensitive camera, to look for several telltale signs of dark energy over a wide swathe of the sky. "This is not the world's biggest telescope, but it has a very large field of view," says Joshua Frieman of the University of Chicago, who is director of the project.

For a start, the telescope will catch many more supernovae. The apparent brightness of each stellar explosion tells us how long ago it happened. During the time the light has taken to reach us, its wavelength has been stretched, or redshifted, by the expansion of space. Put these two things together and we can plot

expansion over time.

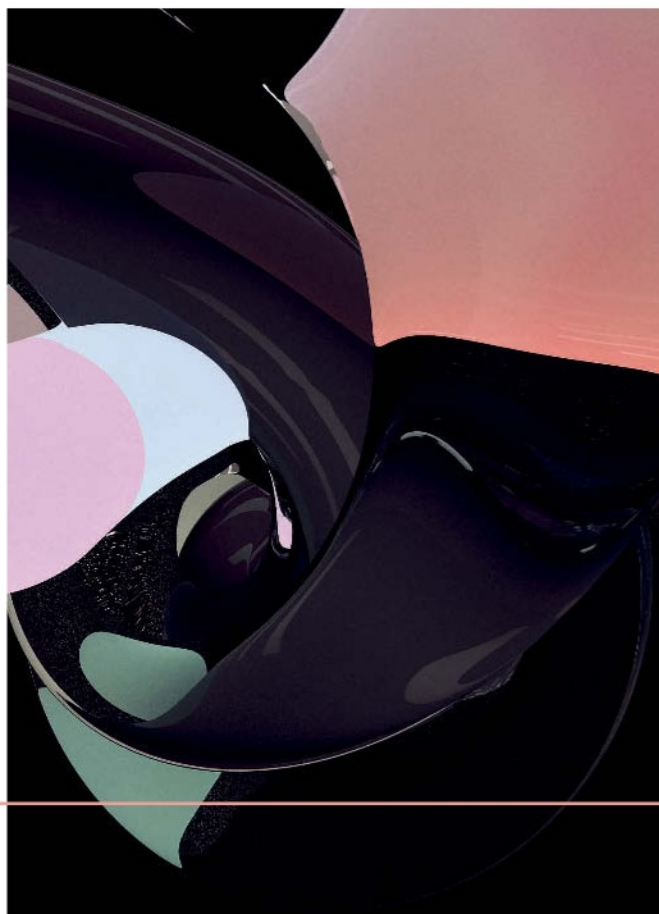
The survey will also draw an intricate sky map that marks the positions of a few hundred million galaxies and their distances from us. Sound waves reverberating around the infant cosmos gave vast superclusters of galaxies a characteristic scale. By measuring the apparent size of superclusters, we can get a new perspective on the expansion history of the universe (see diagram, right).

Eyes on the sky

The map will also reveal dark influences on smaller scales. Dark energy hinders galaxies from coming together to form clusters. The survey team will count clusters directly, and also follow their growth using an effect known as gravitational lensing, which happens when clusters bend light passing through them from even more distant cosmic objects.

These various measurements should give us a handle on how dark energy has changed over time - if at all. The survey should reduce the uncertainty on existing

"If dark energy is a new field or particle, it will create a fifth force we don't feel in the solar system"



DOUGLAS BOWDEN

Stretch the rules

Many lines of evidence indicate that a mysterious **dark energy** is countering gravity's pull and accelerating the universe's expansion

results by a factor of four, says Frieman. With its first analysis due in 2016, it will begin to distinguish between some of the different theoretical models.

A full posse of dark energy hunters will set out just a few years later. The Large Synoptic Survey Telescope, a US-led project, is due to open its great eye in 2021. Other mega-scopes such as the Thirty Meter Telescope in Hawaii and the European Extremely Large Telescope and the Giant Magellan Telescope, both in Chile, should also swing into action around the same time. So should a huge cosmic radio receiver based in Australia and South Africa, the Square Kilometre Array, which will trace cosmic structure through the radio glow of hydrogen clouds. In 2020, the European Space Agency and NASA plan to launch a dark-energy-hunting space mission called Euclid that will trace gravitational lensing and galaxy clumping to even earlier cosmic times. The US Wide-Field Infrared Survey Telescope may follow soon after.

This chase through space will be thrilling, but the quarry may still elude us. Say we find that dark energy maintains a near constant density over time. That would seem to support the cosmological constant, but it would not rule out some quintessence fields that just happen to have a nearly constant density. Even if we find the dark energy density to be increasing or decreasing, we might not be able to tell whether that is due to quintessence, or to some kind of varying gravity.

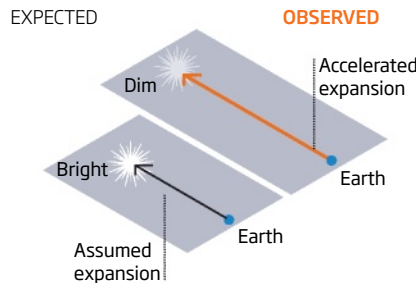
That leads some physicists to suggest laying traps for the beast here on Earth. "If you introduce a new field or particle to be your dark energy, then it will also act as the carrier of a new force," says Clare Burrage at the University of Nottingham, UK. Something like quintessence would produce a fifth fundamental force, separate from gravity, electromagnetism and the nuclear forces. The same holds for most forms of modified gravity. "But we don't see a fifth force within the solar system," says Burrage.

Theorists generally extricate themselves from this sticking point by adding a screening mechanism that weakens the fifth force in comparatively dense environments such as the solar neighbourhood. A project called the GammeV experiment, at Fermilab in Illinois, is already looking for one particular screened dark energy field called the chameleon.

So far GammeV has seen nothing, but now Burrage aims to search for a much wider range of dark energies, and with higher sensitivity. Along with Nottingham colleague Edmund Copeland and Ed Hinds of Imperial College

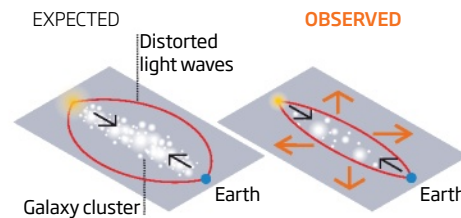
SUPERNOVAE

Distant type Ia supernovae are dimmer than expected, suggesting they are further away



GRAVITATIONAL LENSING

Images of distant galaxies are less distorted by intervening matter than we expected. A repulsive force seems to be stopping matter clumping



London, she wants to expose it using a cloud of cold atoms called a Bose-Einstein condensate, which oscillate together in a collective quantum wave. Dark energy should just slightly slow down the frequency of this oscillation. The team plans to split a condensate in two and place a dense object near one of the halves. If the object screens out dark energy, waves in the two halves will get out of sync, and when brought back together the two condensates will interfere.

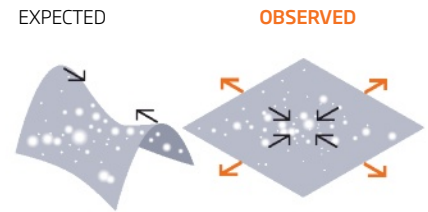
Electric effects

At the University of Washington in Seattle, the Eöt-Wash torsion pendulum experiment is probing other forms of cosmic repulsion. In one theory, extra dimensions of space less than a millimetre across can play host to dark energy. That would also increase the strength of gravity on these scales. A type of screened quintessence called the symmetron would generate a similarly small-scale extra force – a tiny effect that the subtle twistings of the Eöt-Wash pendulums should expose.

Meanwhile, Michael Romalis at Princeton University and Robert Caldwell at Dartmouth College in Hanover, New Hampshire, proposed in 2013 that if ordinary photons or electrons

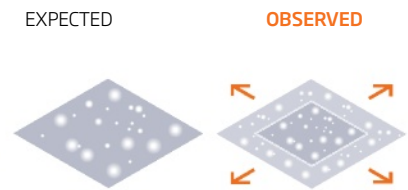
COSMIC MICROWAVE BACKGROUND

With only matter's gravity at work, the universe should be curved. Patterns in the big bang's afterglow suggest it is nearly flat



ACOUSTIC IMPRINTS

Sound waves rippling through the early universe gave galaxy superclusters a typical scale. Far-off superclusters appear smaller than expected – and so are farther away



can feel quintessence even very faintly, then a magnetic field on Earth should generate a tiny electrostatic charge. This effect is potentially simple to detect, although any apparatus designed to do it would need to be extremely precise.

Carroll points out that we might see another electromagnetic effect in space. If photons interact with dark energy, it could rotate their polarisation as they travel across the universe. When the Planck team eventually announce their measurements of the polarisation of photons from the cosmic microwave background, "it is conceivable they will announce they have detected quintessence," he says. We may then have an uneasy wait of a decade or two while the telescopes look out to see which way dark energy is likely to slide, before we can breathe easy that the space around us is not liable to collapse into a new and unhealthy state.

Few imagine that the hunt will be over soon. "Dark energy is one of the greatest mysteries, and I don't expect to still be around when we solve it," says Stephen Hsu of Michigan State University. After 15 years of puzzlement we have no clue as to dark energy's identity. But on the bright side, we do have some clues to where the clues may lie. ■



Pulling power

Dark energy may not be such a mysterious force after all, says [Stephen Battersby](#)



WE WILL be lonely in the late days of the cosmos. Its glittering vastness will slowly fade as countless galaxies retreat beyond the horizon of our vision. Tens of billions of years from now, only a dense huddle of nearby galaxies will be left, gazing out into otherwise blank space.

That gloomy future comes about because space is expanding ever faster, allowing far-off regions to slip across the boundary from which light has time to reach us. We call the author of these woes dark energy, but we are no nearer to discovering its identity. Might the culprit be a repulsive force that emerges

from the energy of empty space, or perhaps a modification of gravity at the largest scales? Each option has its charms, but also profound problems.

But what if that mysterious force making off with the light of the cosmos is an alien echo of light itself? Light is just an expression of the force of electromagnetism, and vast electromagnetic waves of a kind forbidden by conventional physics, with wavelengths trillions of times larger than the observable universe, might explain dark energy's baleful presence. That is the bold notion of two cosmologists who think that such waves could

also account for the mysterious magnetic fields that we see threading through even the emptiest parts of our universe. Smaller versions could be emanating from black holes within our galaxy.

It is almost two decades since we realised that the universe is running away with itself. The discovery came from observations of supernovae that were dimmer, and so further away, than was expected, and earned its discoverers the Nobel prize in physics in 2011.

Prime suspect in the dark-energy mystery is the cosmological constant, an unchanging energy which might emerge from the froth of short-lived, virtual particles that according to quantum theory are fizzing about constantly in otherwise empty space.

Mutant gravity

To cause the cosmic acceleration we see, dark energy would need to have an energy density of about one joule per cubic kilometre of space. When physicists try to tot up the energy of all those virtual particles, however, the answer comes to either exactly zero (which is bad), or something so enormous that empty space would rip all matter to shreds (which is very bad). In this latter case the answer is a staggering 120 orders of magnitude out, making it a shoo-in for the least accurate prediction in all of physics.

This stumbling block has sent some researchers down another path. They argue that in dark energy we are seeing an entirely new side to gravity. At distances of many billions of light years, it might turn from an attractive to a repulsive force.

But it is dangerous to be so cavalier with gravity. Einstein's general theory of relativity describes gravity as the bending of space and time, and predicts the motions of planets and spacecraft in our own solar system with cast-iron accuracy. Try bending the theory to make it fit acceleration on a cosmic scale, and it usually comes unstuck closer to home.

That hasn't stopped many physicists persevering along this route. Until relatively recently, Jose Beltrán and Antonio Maroto were among them. In 2008 at the Complutense University of Madrid, Spain, they were playing with a particular version of a mutant gravity model called a vector-tensor theory, which they had found could mimic dark energy. Then came a sudden realisation. The new theory was supposed to be describing a strange version of gravity, but its equations bore an uncanny resemblance to some of the mathematics underlying another force. ➤

“They looked like electromagnetism,” says Beltrán, now based at the Catholic University of Louvain in Belgium. “We started to think there could be a connection.”

So they decided to see what would happen if their mathematics described not masses and space-time, but magnets and voltages. That meant taking a fresh look at electromagnetism. Like most of nature’s fundamental forces, electromagnetism is best understood as a phenomenon in which things come chopped into little pieces, or quanta. In this case the quanta are photons: massless, chargeless particles carrying fluctuating electric and magnetic fields that point at right angles to their direction of motion.

Alien photons

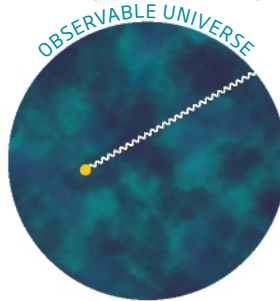
This description, called quantum electrodynamics or QED, can explain a vast range of phenomena, from the behaviour of light to the forces that bind molecules together. QED has arguably been tested more precisely than any other physical theory, but it has a dark secret. It wants to spit out not only photons, but also two other, alien entities.

The first kind is a wave in which the electric field points along the direction of motion, rather than at right angles as it does with ordinary photons. This longitudinal mode moves rather like a sound wave in air. The second kind, called a temporal mode, has no magnetic field. Instead, it is a wave of pure electric potential, or voltage. Like all quantum entities, these waves come in particle packets, forming two new kinds of photon.

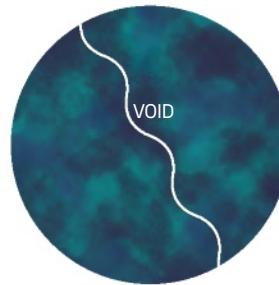
As we have never actually seen either of these alien photons in reality, physicists found a way to hide them. They are spirited away using a mathematical fix called the Lorenz condition, which means that all their

Elongated light

The theory of dark magnetism suggests that a period of inflation in the early universe could have unleashed modes of light whose wavelengths are simply too big for them to be observed



NORMAL PHOTON
Observable wavelengths from trillionths of a metre to billions of metres



LONGITUDINAL MODE
Longer wavelengths, but still less than observable universe. Their presence could explain magnetic fields in voids



TEMPORAL MODE
Wavelengths larger than observable universe. Their energy could account for dark energy and the accelerating expansion of the universe

attributes are always equal and opposite, cancelling each other out exactly. “They are there, but you cannot see them,” says Beltrán.

Beltrán and Maroto’s theory looked like electromagnetism, but without the Lorenz condition. So they worked through their equations to see what cosmological implications that might have.

The strange waves normally banished by the Lorenz condition may come into being as brief quantum fluctuations – virtual waves in the vacuum – and then disappear again. In the early moments of the universe, however, there is thought to have been an episode of violent expansion called inflation, which was driven by very powerful repulsive gravity. The force of this expansion grabbed all kinds of quantum fluctuations and amplified them hugely. It created ripples in the density of matter, for example, which eventually seeded galaxies and other structures in the universe.

Crucially, inflation could also have boosted the new electromagnetic waves. Beltrán and Maroto found that this process would leave behind vast temporal modes: waves of electric potential with wavelengths many orders of magnitude larger than the observable universe. These waves contain some energy but because they are so vast we do not perceive them as waves at all. So their energy would be invisible, dark... perhaps, dark energy?

Beltrán and Maroto called their idea dark magnetism. Unlike the cosmological constant, it may be able to explain the actual quantity of dark energy in the universe. The energy in those temporal modes depends on the exact time inflation started. One plausible moment is about 10 trillionths of a second after the big bang, when the universe cooled below a critical temperature and electromagnetism split from the weak nuclear force to become a force in its own right. Physics would have suffered a sudden wrench, enough perhaps to provide the impetus for inflation.

If inflation did happen at this “electroweak transition”, Beltrán and Maroto calculate that it would have produced temporal modes with an energy density close to that of dark energy. The correspondence is only within an order of magnitude, which may not seem all that precise. In comparison with the cosmological constant, however, it is mildly miraculous.

The theory might also explain the mysterious existence of large-scale cosmic magnetic fields. Within galaxies we see the unmistakable mark of magnetic fields as they twist the polarisation of light. Although the turbulent formation and growth of galaxies could boost a pre-existing field, is it not clear where that seed field would have come from.

“The equations for this strange gravity bear an uncanny resemblance to those of another force: electromagnetism”



Even more strangely, magnetic fields seem to have infiltrated the emptiest deserts of the cosmos. Their influence was noticed in 2010 by Andrii Neronov and Ievgen Vovk at the Geneva Observatory, Switzerland. Some distant galaxies emit blistering gamma rays with energies in the teraelectronvolt range. These hugely energetic photons should smack into background starlight on their way to us, creating electrons and positrons that in turn will boost other photons up to gamma energies of around 100 gigaelectronvolts. The trouble is that astronomers see relatively little of this secondary radiation. Neronov and Vovk suggest that is because a diffuse magnetic field is randomly bending the path of electrons and positrons, making their emission more diffuse.

"It is difficult to explain cosmic magnetic fields on the largest scales by conventional mechanisms," says astrophysicist Larry Widrow of Queen's University in Kingston, Ontario, Canada. "Their existence in the voids might signal an exotic mechanism." One suggestion is that giant flaws in space-time called cosmic strings are whipping them up.

With dark magnetism, such a stringy solution would be superfluous. As well as the gigantic temporal modes, dark magnetism should also lead to smaller longitudinal waves bouncing around the cosmos. These waves could generate magnetism on the largest scales and in the emptiest voids.

To begin with, Beltrán and Maroto had some qualms. "It is always dangerous to modify a well-established theory," says Beltrán. Cosmologist Sean Carroll at the California Institute of Technology in Pasadena, echoes this concern. "They are doing extreme violence to electromagnetism. There are all sorts of dangers that things might go wrong," he says. Such meddling could easily throw up absurdities, predicting that electromagnetic forces are different from what we actually see.

The duo soon reassured themselves, however. Although the theory means that temporal and longitudinal modes can make themselves felt, the only thing that can generate them is an ultra-strong gravitational field such as the repulsive field that sprang up in the era of inflation. So within the atom, in all our lab experiments, and out there among the planets, electromagnetism carries on in just the same way as QED predicts.

Carroll is not convinced. "It seems like a long shot," he says. But others are being won over. Gonzalo Olmo, a cosmologist at the University of Valencia, Spain, was initially sceptical but is now keen. "The idea is fantastic. If we quantise

"The theory might explain the mysterious appearance of galactic-scale magnetic fields in places they ought not to be"

electromagnetic fields in an expanding universe, the effect follows naturally."

So how might we tell whether the idea is correct? Dark magnetism is not that easy to test. It is almost unchanging, and would stretch space in almost exactly the same way as a cosmological constant, so we can't tell the two ideas apart simply by watching how cosmic acceleration has changed over time.

Ancient mark

Instead, the theory might be challenged by peering deep into the cosmic microwave background, a sea of radiation emitted when the universe was less than 400,000 years old. Imprinted on this radiation are the original ripples of matter density caused by inflation, and it may bear another ancient mark. The turmoil of inflation should have energised gravitational waves, travelling warps in space-time that stretch and squeeze everything they pass through. These waves should affect the polarisation of cosmic microwaves in a distinctive way, which could tell us about the timing and the violence of inflation. The European Space Agency's Planck spacecraft is looking for this signature, and observations at the BICEP2 telescope based at the South Pole suggest the waves are indeed there. If astronomers find that inflation happened before the electroweak transition, at a higher energy scale, then that would rule out dark magnetism in its current form.

Olmo thinks that the theory might anyhow need some numerical tweaking, so that might not be fatal, although it would be a blow to lose the link between the electroweak transition and the correct amount of dark energy.

One day, we might even be able to see the twisted light of dark magnetism. In its present incarnation with inflation at the electroweak

scale, the longitudinal waves would all have wavelengths greater than a few hundred million kilometres, longer than the distance from Earth to the sun. Detecting a light wave efficiently requires an instrument not much smaller than the wavelength, but in the distant future it might just be possible to pick up such waves using space-based radio telescopes linked up across the solar system. If inflation kicked in earlier at an even higher energy, as suggested by Olmo, some of the longitudinal waves could be much shorter. That would bring them within reach of Earth-based technology. Beltrán suggests that they might be detected with the Square Kilometre Array – a massive radio instrument due to come on stream within the next decade.

If these dark electromagnetic waves can be created by strong gravitational fields, then they could also be produced by the strongest fields in the cosmos today, those generated around black holes. Beltrán suggests that waves may be emitted by the black hole at the centre of the Milky Way. They might be short enough for us to see – but they could easily be invisibly faint. Beltrán and Maroto are looking into it.

One thing they have calculated from their theory is the voltage of the universe. The voltage of the vast temporal waves of electric potential started at zero when they were first created at the time of inflation, and ramped up steadily. Today, it has reached a pretty lively 10^{27} volts, or a billion billion gigavolts.

Just as well for us that it has nowhere to discharge. Unless, that is, some other strange quirk of cosmology brings a parallel universe nearby. The encounter would probably destroy the universe as we know it, but at least then our otherwise dark and lonely future would end with the mother of all lightning bolts. ■



CHAPTER FIVE

BLACK HOLES

THEY are the great dark spots of the universe. Not merely black holes, but enormous black holes, billions of times the mass of our sun. They are everywhere we look, even looming distantly out of the light of the cosmic dawn.

Supermassive black holes have all the space-warping strangeness of their smaller kin, which we see dotted about our galaxy; they too swallow all matter, light and cries for help. But they hold another level of mystery. We know that small black holes with a mass a few times that of the sun are born when a large star's heart collapses in a supernova, but nobody can explain the genesis of the giants.

We thought we could. We thought supermassive black holes simply grew slowly from very small seeds, bloating themselves on surrounding gas. But our observations show this cannot be the whole story. We need a new explanation for these ancient monsters.

Whatever did make supermassive black holes, it had to make an army. Observations of stars whirled around by powerful gravity suggest there is a huge black hole at the heart of almost all large galaxies, including our own. The Milky Way's is 5 million times the mass of the sun. Wander some 50 million light years

away to the giant elliptical galaxy M87, and you will find a black hole of over 6 billion solar masses. Its event horizon – the boundary of no return for any passing object – is nearly five times as wide as the orbit of Neptune.

Further off we see even more dramatic signs of the supermassive. Extraordinarily active galaxies called quasars have a brilliant point of light at their heart that often outshines all the

“There is a limit to how fast a black hole can grow: eventually it is starved by its own brilliant belching”

billions of surrounding stars together. Many spit out searing flares of X-rays and gamma rays and colossal jets of material at 99 per cent of the speed of light. These are the signs of a giant black hole's voracious appetite. Gas spiralling inwards heats through friction and glows white-hot, generating magnetic fields that send the jets of matter hurtling outwards.

The conventional tale of supermassive black holes begins a few tens of millions of years after

the big bang, as the very first stars formed from the densest clouds of primordial hydrogen and helium gas. These pioneers, the story goes, were several hundred times the mass of our sun. The core of such a star soon collapses to form a black hole of around 100 solar masses. As this seed gorges on gas it sinks towards the centre of its galaxy, eventually becoming the powerful heart of a quasar.

In 2000, however, NASA's Chandra X-ray telescope spied a very distant and powerful quasar. We see SDSS J1030+0524 as it was just 900 million years after the big bang, and its power output must come from a hole of more than a billion solar masses.

How did this monster get so vast so fast? The more gas a black hole gulps, the more light and other radiation shines out. Eventually, the hole is starved by its own brilliant belching: the flood of light grows so fierce that it sweeps away any incoming gas, cutting off the food supply. Theory, backed by the behaviour of black holes nearby, says that a hole can double in size at most once every 30 million years.

To make the leap from 100 to a billion solar masses means doubling in mass 23 times, so in theory the black hole in SDSS J1030+0524 could have developed in around 700 million years. But that would require a gas supply perfectly tuned to the hole's needs. A black hole's surroundings are likely to be messy and changeable, so gorging so fast for so long is not so easy. “It is a problem to imagine they can accrete gas at all times,” says theorist Zoltan Haiman of Columbia University in New York.

Still, SDSS J1030+0524 might be a freakish example of a black hole managing to live the high life for the best part of a billion years. “One peculiar object you can always explain,” says astrophysicist Priya Natarajan of Yale University. But we have discovered scores more at comparable distances. “When there's a population of these things, there has to be a natural way to make them,” says Natarajan.

Each discovery piles on the pressure. In 2012, a team using the UK Infrared Telescope in Hawaii observed a quasar, ULAS J1120+0641, with a mass about 2 billion times that of the sun just 770 million years after the big bang. Theory dictates a minimum of about 750 million years to grow that large from a 100-solar-mass start.

The story of supersized holes growing from small seeds looks even less plausible after recent work on how the first stars formed. ➤

Monster mix

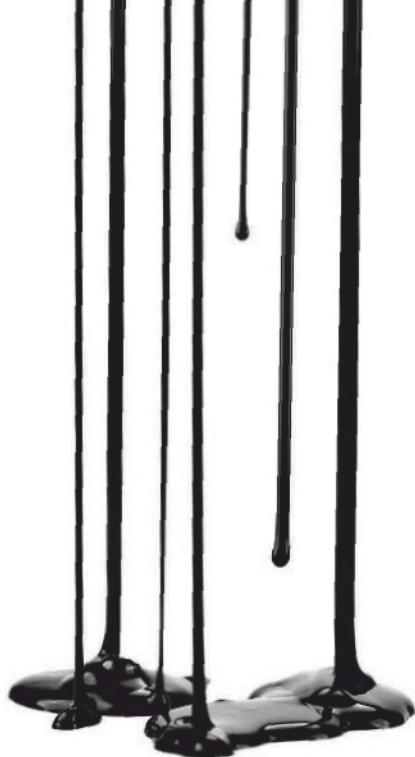
Giant black holes pop up even in the deep recesses of time. How did they get so vast so fast, asks **Stephen Battersby**

New simulations have followed the shrinking gas clouds that gave birth to them for a longer spell, and found that they tend to split up into smaller fragments than we thought, making stars no larger than about 50 solar masses. After exploding, these would make black holes of only about 10 solar masses – a disappointing size for would-be quasar seeds. “They are just not oomphy enough,” says Natarajan.

What’s more, common-or-garden stellar-mass black holes should pop up all over every young galaxy. Some of these holes would sink to the centre of their galaxy and seed more massive ones, so even fairly small galaxies should sport a fairly impressive central black hole by now. But that is not what we see. Jenny Greene of Princeton University has found that among smallish galaxies with a total mass of about a billion suns, only about half have a central hole.

There seems to be only one conclusion. We’re going to need a bigger seed.

One possibility is that supermassive black holes began not with single stars, but many. “We know that early in the history of the universe, stars tended to form in bursts – regions that were spectacularly active,” says Fred Rasio of Northwestern University in Evanston, Illinois. In 2003, he ran simulations of ancient clusters where hundreds of bright young things were forming. The most massive tended to pile up near the centre, where they almost inevitably run into one another. “You form a thing – I don’t want to call it a star – with many thousands of solar masses,” he says. What happens next is very difficult to model. Rasio’s educated guess is that this object might well collapse to form a black hole of perhaps a few thousand solar masses.



“Perhaps the most radical suggestion is that giant black holes were forged directly in the big bang”

It’s a nice idea. It would be nicer still if we could find similar middleweight black holes in star clusters today. A few promising objects called ultraluminous X-ray sources (ULXs) have been found in nearby galaxies. They are seemingly bright enough to be based around bigish holes. But in 2011, observations of one ULX in our neighbouring galaxy Andromeda showed that it has the same characteristic spectrum and behaviour as small black holes in our galaxy, with around 10 solar masses.

The other ULXs may be this small too.

In any case, even a seed hole of a 1000 solar masses would have to double its mass 20 times over to become a gigasun giant. Almost constant gorging would still be needed to explain an object like ULAS J1120+0641.

Perhaps, then, we need to think a little bigger still. If a small black hole springs from the collapsing heart of a star, might a big one come from the collapsing heart of a galaxy?

This was originally suggested as an outside possibility in 1978 by Martin Rees at the University of Cambridge. It sounds seductively simple, but it is not easy to cram so much matter into a galaxy’s heart.

The first hurdle is spin. Even the earliest protogalaxies rotated a little, tweaked by their neighbours’ gravity. As they contracted, their gas whirled faster, like air drawn into a tornado. Eventually, the rotation balanced gravity, producing a spinning disk of gas with little material within the innermost few hundred light years.

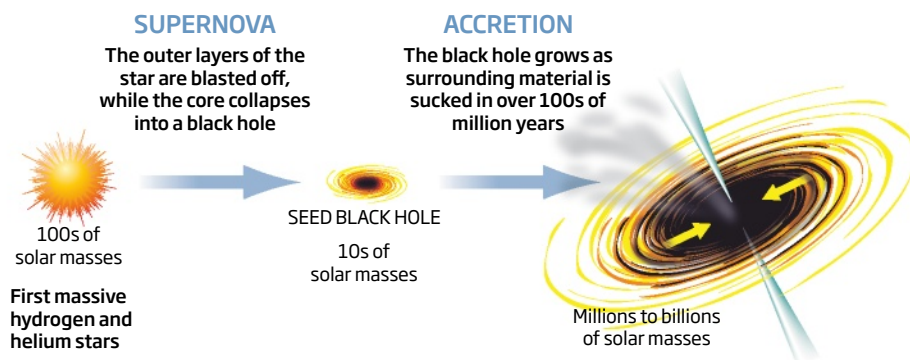
Rasio and Abraham Loeb of Harvard University showed in the 1990s how this barrier might be overcome. If a protogalaxy is slow-spinning and dense, its core can become unstable. Excess gas collects into rotating, elongated bars, which act like gravitational gear-wheels transferring the rotation outwards. The heart of the protogalaxy can then collapse into a much denser knot.

The next stage is uncertain, but according to calculations performed in 2006 by Rees, his then colleague Marta Volonteri and Mitchell Begelman of the University of Colorado at Boulder, one possibility is a monstrous “quasistar”: a dense cocoon of gas a few hundred million kilometres across surrounding a small central black hole. The cocoon’s great weight would force matter into the hole, incubating it to a million solar masses in just a few tens of millions of years. Such a heavy seed could square the theory with observations: it would need to double in size just 10 times to make a billion-sun black hole. Even without a carefully balanced gas supply, that could happen within the allotted time frame of about 700 million years.

Problems remain. During its initial collapse, the gas is liable to split up into little blobs that will coalesce and ignite into stars, denying material to the quasistar. Some of these stars will go supernova, blasting away fresh gas supplies and halting the black hole’s growth. There are ways around this: ultraviolet radiation from nearby starbursts could heat the blobs and stop them coalescing, or turbulence might prevent fragmentation.

Making a big one

The old picture of supermassive black hole formation has the cosmic giants feeding constantly on surrounding gas over hundreds of millions of years



To many astronomers, however, these remedies seem too contrived. "Direct collapse needs a lot of fine-tuning," says Rasio.

Do these failures mean something altogether stranger is going on? Perhaps the most radical suggestion is that giant black holes were forged directly in the fires of the big bang, during tumultuous moments known as phase transitions when matter and radiation suddenly rearranged themselves. About one microsecond after the beginning of time, for example, quarks were coming together to form protons and neutrons. This process could have been uneven, producing

sharp spikes in density that turned into black holes of around one solar mass.

That is too small to make our seeds, but Sergei Rubin at the National Research Nuclear University in Moscow, Russia, has suggested that these holes might cluster together and swiftly merge into one giant. Another promising phase transition happens when the universe is about 10 seconds old, when a haze of electrons and positrons destroy one another to leave space filled instead with gamma ray photons. At this point, holes of up to 100,000 solar masses might form spontaneously.

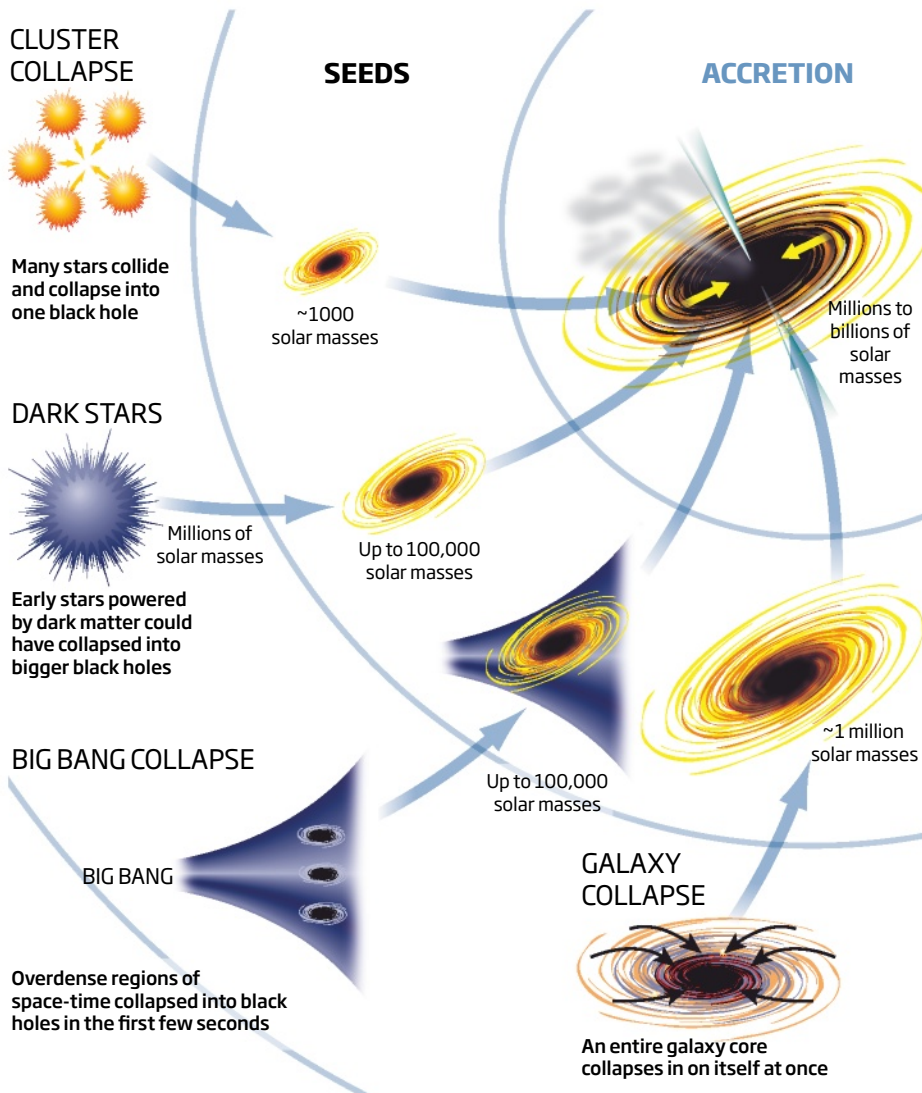
These mammoths of the universal dawn would suck in hot gas around them and shine brightly in X-rays, leaving marks on the cosmic microwave background radiation. Searches have so far found nothing, but do not rule out a few rare giant primordial holes, enough to form the seeds of early quasars.

Even so, most astronomers trying to fathom the origin of supermassive holes think that primordial seeds are an unnecessarily exotic option. "There is no compelling argument for why they should have formed," says Begelman. "You need some weird cosmology."

Perhaps more palatable is the notion of dark stars. In 2007, Douglas Spolyar, then at the University of California, Santa Cruz, and colleagues suggested that some of the very first stars might have been powered by dark matter. Dark matter particles would settle into the stellar core, where they could collide with and annihilate one another, providing a gentler heat than the nuclear furnaces of hydrogen and helium-powered giants. With no harsh radiation to blow away incoming gas, dark stars could keep growing almost indefinitely, finally collapsing to form a seed black hole of up to 100,000 solar masses.

Growth spurt

Quasar observations suggest supermassive black holes already existed in the early universe - meaning we need to explain how their seeds got big so quick



Point of no return

This idea is very testable. Dark stars would emit enough infrared radiation to be seen by the successor of Hubble, the James Webb Space Telescope, which is scheduled for launch in 2018. It might also be able to detect quasistars.

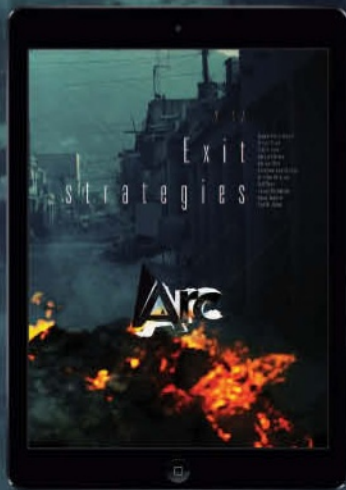
If it sees nothing, we will need an even more sophisticated piece of kit to spy on the origin of supermassive black holes. The Laser Interferometer Space Antenna is designed to detect gravitational waves, travelling space warps that relativity says should be produced copiously by colliding black holes. LISA has spent a long time trying to get off the drawing board, but if it finally blasts into space it could detect the gravitational waves from merging holes across the universe. "Then it should be easy to tell whether the seeds are small or massive," says Volonteri.

Meanwhile, astronomers will be searching for quasars yet deeper in space and time. What happens if they keep finding them ever earlier in the history of the universe? Eventually, they will reach a point a couple of hundred million years after the big bang when their black centres would not have had time to grow even from a million-sun seed. Then none of these stories would work. We would truly be in the dark about why the universe is full of holes. ■



Exit strategies

The newest issue of Arc
is available now



Arc uses stories, features, essays and provocations to map the future. In this issue we ask: is there a back door to the universe? And if there is, who's hidden the key?

Featured writers in this issue with new science fiction, essays and ideas about the future include:

"Arc lays out our human choices through fact-informed fiction and imagination-inspired fact, test-driving our hopes and fears."

— Margaret Atwood

Jeff Noon
Kathleen Ann Goonan
Tad Williams
Hannu Rajaniemi
M. John Harrison
Claire Dean
and more over 150 pages

Don't miss out on *Exit strategies* - buy it now in a range of digital formats.
Find out more at arcfinity.org

A BOMB made out of a black hole is a rather unsettling thought. The idea of a monstrous entity consuming everything that ventures too close is sinister enough. You wouldn't want one exploding in front of you in a searing flash of heat and light.

Yet according to two astrophysicists in the US, something like that might actually have happened – not in today's cosmos, but way back in the fireball of the big bang. In fact, laser-powered black hole bombs may have been going off like firecrackers across the length and breadth of our infant universe.

If so, they would give us a valuable new window on the exotic physics that went on in our universe's obscure first moments, and perhaps also help to explain one of its biggest contemporary mysteries – where most of its mass is hiding. A crazy idea? Perhaps – but the best thing is, we can test it right now.

We know that two distinct classes of black hole exist. There are stellar-mass ones, weighing in at anything between a few times and many tens of times the mass of the sun.

These form when very massive stars collapse, and are thought to pockmark galaxies including our own, the Milky Way. Then there are “supermassive” black holes, weighing in at anything up to 30 billion solar masses. These are thought to form in the heart of most galaxies – again, including our own – when a dense star cluster or gas cloud collapses into a black hole and begins to pull in, or accrete, more material from its surroundings.

What unites both sorts of black hole is an agglomeration of mass so great in so small a space – a “singularity” in space-time – that nothing can escape its gravitational pull. Everything that gets too close is simply sucked in, never to be seen again.

The idea that one of these monsters might make a bomb dates back four decades. In 1974, theorists William Press and Saul Teukolsky noted that if a black hole were spinning fast enough, light of a long-enough wavelength passing close by would scatter off it, rather than being sucked in. If this spinning black hole were to be surrounded by something ➤

Black bombshells

Was the early universe riven by monstrous laser-powered explosions, asks **Marcus Chown**

MARK GREINER/SHUTTERSTOCK

like a mirror, the light could be reflected and scattered many times. Fuelled by energy from the rotation of the black hole, it would bounce back and forth and amplify itself in a runaway process rather like what happens in the mirrored cavity within a laser. If the surrounding mirror were removed or shattered, the light would instantaneously escape in a powerful burst of light and heat – a black hole bomb.

Little big crunch

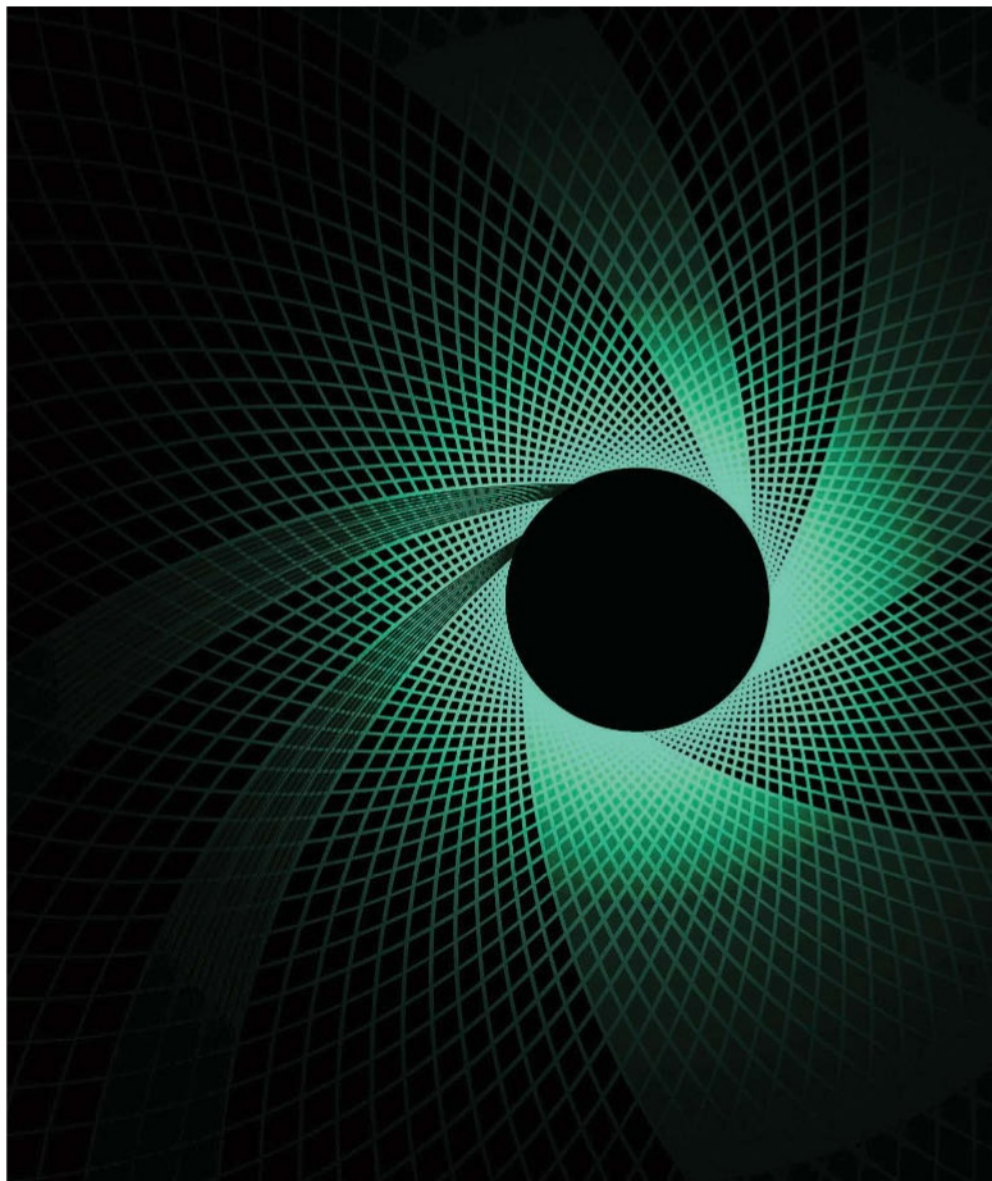
It sounds like a theorist's pipe dream, but then physics is a strange business. Actually, exactly this sort of set-up could have existed in the universe's infancy, says Abraham Loeb, who is head of astronomy at Harvard University. That is, it could if a third class of black hole were to exist beyond the stellar-mass and supermassive types.

The basic premise of such "primordial" black holes is simple, and the idea has been around for a while. Our observable universe is surrounded by a horizon that marks the furthest points from which light can have travelled to us in the 13.8 billion years or so since the big bang. If the density of matter within this horizon is greater than a critical value, its gravitational attraction will at some point in the future cause everything to collapse down on itself in a "big crunch" – a kind of big bang in reverse.

The same would have been true of the universe in its earliest moments. Back then, the horizon seen from any point in space would have encompassed a much smaller region. If any particular region within the wider universe were to have a density greater than the critical density, it would begin to collapse. "All the stuff within it – essentially photons – would shrink down in a big crunch to form a primordial black hole," says Loeb.

As the universe grows, the potentially collapsible regions within it grow too. Because these regions contain more mass, the mass of the black holes that can form in this way also increases. At the same time, though, the likelihood of sustaining a large enough density fluctuation to start such a collapse decreases. This means that if primordial black holes exist, the bulk of them are likely to be very small; from microscopic ones with masses less than that of the moon to ones the size of a fridge with the mass of Jupiter.

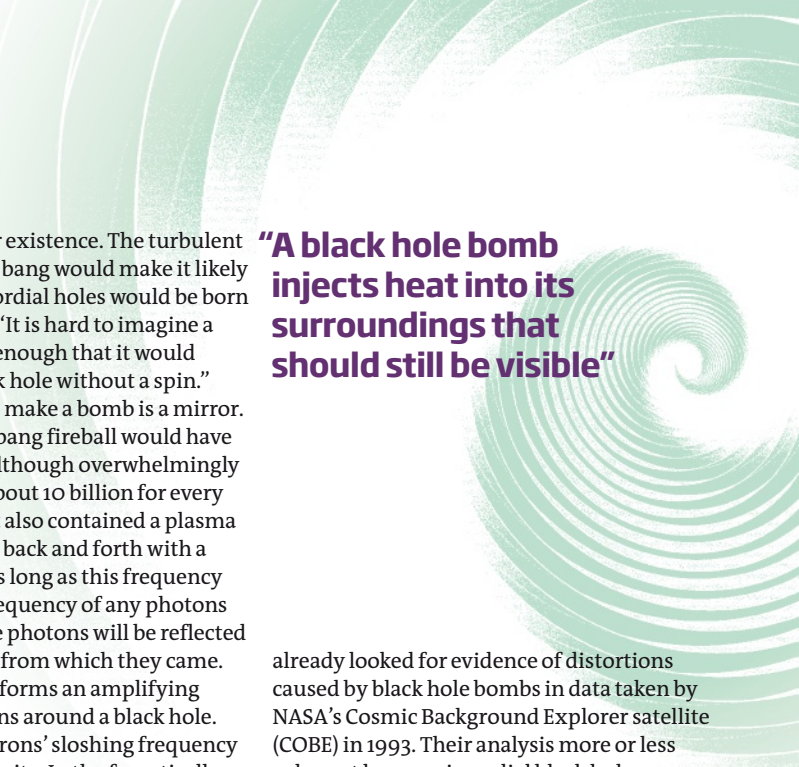
Loeb thinks such black holes might provide an identity for dark matter. To make our standard cosmology work and keep galaxies and the like stuck together, this invisible stuff must outweigh visible stars and galaxies by a



"If primordial black holes exist, they will be very small - no bigger than a fridge"

factor of about 5.5. Most physicists believe it takes the form of a soup of hitherto-unknown subatomic particles, but Loeb suggests they are barking up the wrong tree. "The subatomic particles are predicted by speculative theories of particle physics, whereas we have strong observational evidence that black holes exist, spanning a wide range of masses," he says. "It's really not such a leap of faith to imagine them as the dark matter."

Experiments have already put some limits on how much primordial black holes of different sizes can contribute to dark matter. Dark objects passing in front of stars would boost, or "microlens", the starlight as it travels towards us, making the star appear briefly brighter. Various probes have looked for this effect and seen very little evidence for it. Primordial black holes in the range from about



behind proof of their existence. The turbulent conditions of the big bang would make it likely that these tiny primordial holes would be born spinning, says Loeb. "It is hard to imagine a collapse symmetric enough that it would have spawned a black hole without a spin."

Now all we need to make a bomb is a mirror. No problem, the big bang fireball would have provided that, too. Although overwhelmingly made of photons – about 10 billion for every particle of matter – it also contained a plasma of electrons sloshing back and forth with a natural frequency. As long as this frequency is greater than the frequency of any photons that encounter it, the photons will be reflected back in the direction from which they came. The electron plasma forms an amplifying mirror for the photons around a black hole.

Crucially, the electrons' sloshing frequency depends on their density. In the frenetically expanding early universe, this density was constantly falling. At some point, the electron frequency would have fallen below the critical threshold needed to contain the photons. They would have been released in searingly powerful bursts of heat and light. Voila – black hole bombs.

Visible signatures

That has observational consequences. "Crucially, a black hole bomb injects heat into its immediate surroundings," says Loeb. With his colleague Paolo Pani of the Technical University of Lisbon, Portugal, he has shown that this heat should still be visible as deviations from the perfect "black body" spectrum expected of the cosmic microwave background, the relic radiation of the big bang that suffuses the cosmos today at a frigid temperature of 2.7 kelvin.

It is an idea worth investigating, says John Mather, a physicist at the NASA Goddard Space Flight Center in Greenbelt, Maryland, and co-winner of the 2006 Nobel prize in physics for his investigations of the cosmic microwave background. "I have no idea why nobody noticed it before. It's pretty interesting," he says.

Fabio Capela at the Free University of Brussels in Belgium agrees. "This is an important piece of work. Such a situation has not been considered in any previous analysis," he says. But he sounds a note of caution. "All of the work is based on the assumption that primordial black holes acquire a spin when they are created," he says. "Nobody knows if this is true."

So let's take a look at the cosmic background radiation to find out. Loeb and Pani have

"A black hole bomb injects heat into its surroundings that should still be visible"

already looked for evidence of distortions caused by black hole bombs in data taken by NASA's Cosmic Background Explorer satellite (COBE) in 1993. Their analysis more or less rules out larger primordial black holes making up more than about 1 per cent of the missing dark-matter mass today, but still leave open the possibility that smaller ones might make up all of it. Experiments are on the drawing board to study the cosmic microwave background in even finer detail. They should either see something, or narrow things down still further.

The interest in black hole bombs is not just confined to dark matter. Observations tell us that stellar-mass and supermassive black holes can also spin. Theoretically, they too could kick off a lasing effect – not with photons, but with other, so far undiscovered particles similar to photons that might not have anything to do with dark matter. Any evidence of such bombs – or lack of it – could put stringent limits on the existence of such particles. "We haven't done the test yet, but it is eminently possible," says Loeb.

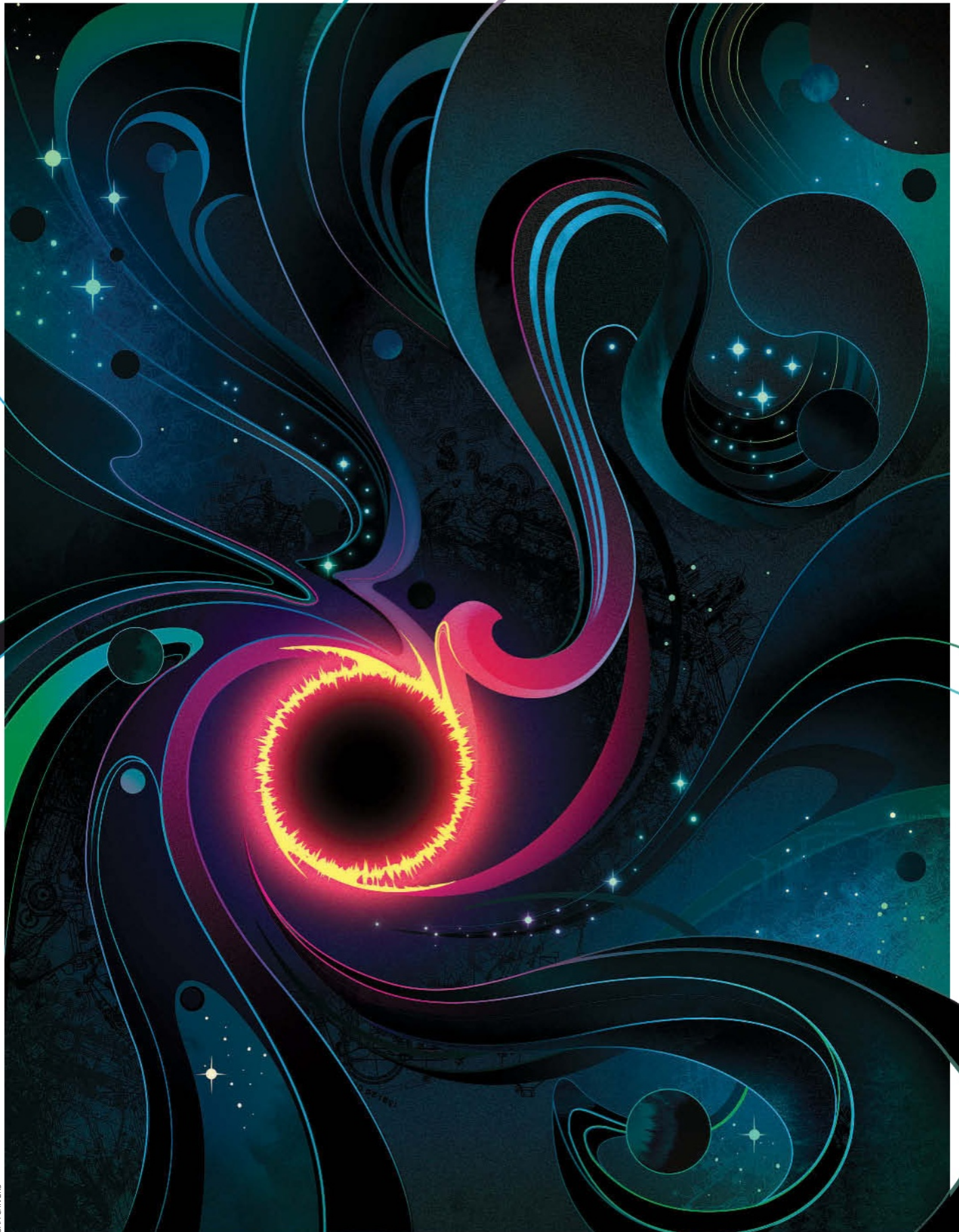
But it is undoubtedly those smallest primordial black holes, potentially forged in the earliest moments of the big bang, that would be the most exciting catch. One process that could have caused the density within a given space in the early universe to be higher than average, and so have created them, is a change of state akin to the transition from steam to liquid water that liberates a lot of heat energy. Evidence for such changes of state is eagerly sought, as they would take us beyond the standard model of particle physics to a new and deeper level of understanding. "Evidence of primordial black holes at a particular mass will tell us about physical processes operating at a very particular time," says Loeb. "It could provide a unique window on exotic events going on in the first moments of the universe's existence."

Now that really would be a bombshell. ■

one-millionth the mass of the sun – about a third the mass of Earth's moon – up to 10 solar masses can make up 10 per cent of dark matter at most.

A further limit comes from the prediction that, over time, black holes evaporate away, giving out so-called Hawking radiation. Primordial black holes with less than the mass of a typical asteroid, about one-hundred-millionth the mass of the sun, would have evaporated by now. Since astronomers see no sign of the flashes of high-energy gamma rays expected to accompany this process, these holes can be ruled out as a big component of dark matter.

That still leaves unexplored a range of masses for primordial black holes between asteroid and moon size. Some of these might have exploded as bombs – and so have left



Ring of fire

Falling into a black hole has never been more dangerous. **Anil Ananthaswamy** reports from a safe distance

PARADOXES are good in physics,” reflects John Preskill. “They help to point the way towards important discoveries.” Quantum mechanics and Einstein’s theories of relativity offer plenty to choose from. There’s the cat that can be dead and alive at the same time. Or the *Back to the Future*-style time traveller who kills his own grandfather, rendering his own birth impossible. Or the twins who disagree on their age after one returns from a near light-speed trip to a neighbouring star. Each perplexing scenario forces us to examine the fine print of the problem, thereby advancing our understanding of the theory behind it. A case in point is Einstein, whose own theories came from trying to resolve the paradoxes of his time.

Now Preskill, a theoretical physicist at the California Institute of Technology in Pasadena, is scratching his head over the latest one to surface. Nicknamed the black hole firewall paradox, it comes about when you consider what happens to someone falling into a black hole.

With the nearest black hole more than 1000 light years away, the question is very much a theoretical one. Yet just by studying such a possibility, physicists are hoping to make a breakthrough in their efforts to combine general relativity and quantum mechanics into a theory of quantum gravity – one of the most intractable problems in physics today.

Black holes have long been fertile breeding grounds for paradoxes. Back in 1974, Stephen Hawking, along with Jacob Bekenstein of the Hebrew University in Jerusalem, Israel, famously showed that black holes are not

entirely black. Instead, they radiate energy known as Hawking radiation comprising photons and other quantum particles – an agonisingly slow process that eventually causes the black hole to evaporate completely.

Hawking spotted a problem with this picture. The radiation seemed so random that he surmised it couldn’t carry any information about the stuff that had fallen in. So as the black hole evaporates, the information it holds must eventually disappear. Yet this is in direct conflict with a central tenet of quantum physics, which says that information cannot be destroyed. The black hole information paradox was born.

Over the decades, physicists have struggled with this paradox. Hawking thought that black holes destroyed information and the answer was to question quantum mechanics. Others disagreed. After all, Hawking’s idea came from his efforts to meld general relativity and quantum mechanics – a mathematical feat so elusive that he was forced to make approximations. Preskill even made a bet with Hawking that black holes don’t destroy information.

Several arguments suggest that Hawking was wrong. One of the most compelling comes from thinking about what happens as the evaporating black hole gets smaller and smaller. If information can’t escape or be destroyed, then more and more has to be stored in an ever-shrinking volume. But if this is the case, quantum theory says the probability for making a tiny black hole increases from virtually nothing to almost infinity wherever matter collides against matter. “You should have seen it at the Large Hadron Collider, ➤

"You're sailing along through space-time and then, bam! You hit this firewall and burn up"

you should have seen it at Fermilab, you should have seen it in tiny room-sized particle accelerators from the 1930s," says Don Marolf, a theorist at the University of California in Santa Barbara (UCSB). "You should see it when you go and jump up and down on the grass."

Obviously that hasn't happened. The other possibility – that matter and the information it carries can leak out from a black hole – is unlikely. Any material that falls in would need to travel faster than light to escape the black hole's fearsome gravity.

Perhaps, instead, the answer lies with the Hawking radiation itself. Maybe it isn't so featureless. "A common reaction was that Hawking had simply been careless," says Joseph Polchinski, also at UCSB. "It wasn't that information was lost, it was that he hadn't kept track of it enough."

Yet all early efforts to do away with the paradox proved unsuccessful. "Hawking had identified a really deep problem," says Polchinski.

As it happened, Hawking changed his mind in 2004, partly due to work by an Argentinian physicist called Juan Maldacena (see "Hawking's change of heart", right). Black holes don't destroy information after all, he conceded. He honoured the bet he made with Preskill and presented him with an encyclopaedia of baseball, which Preskill likened to a black hole, because it was heavy and it took effort to get information out of it.

Into the abyss

Attention soon shifted to wondering how information could get out of a black hole. This isn't an easy question to answer. And it is by exploring these issues that the new black hole firewall paradox has come into sharp focus.

A firewall is the catchy new term for something physicists have long suspected might happen if information escapes from a black hole. To understand the argument, we need a simplified picture of Hawking radiation. The vacuum of space-time is constantly producing pairs of virtual particles, which pop into existence and just as quickly disappear. This picture changes near the black hole's event horizon, considered the point of no return for anything falling in. Occasionally one of the pair is sucked into the black hole while the other escapes. It is a rare particle

that flees a black hole, but it is these that constitute Hawking radiation.

Now if Hawking radiation is carrying quantum information out, then that creates a problem. One of Hawking's great insights was to show how quantum theory, general relativity and thermodynamics are all linked at a black hole. This means that the paired particles just inside the event horizon would become immensely energetic as information is transferred to their partners outside, creating a wall of fire hot enough to burn up anything, or anyone, falling into the black hole.

This dramatically contradicts anything general relativity tells us about black holes. In fact, such firewalls seemed so preposterous that physicists started looking for ways to transfer information out of the black hole without such transgressions.

One possibility has been put forward by Steve Giddings, also at UCSB. Building on work done by Samir Mathur of Ohio State University, Giddings developed a simple model of a black hole. The work showed that

HAWKING'S CHANGE OF HEART

It was string theorist Juan Maldacena who made the breakthrough that eventually led to Stephen Hawking changing his mind about black holes and information (see main story). In 1997, Maldacena used the mathematics of string theory to show that the theory of gravity that describes the insides of a black hole is equivalent to the quantum theory that describes the black hole's surface. It sounds esoteric, yet Maldacena's work is remarkable. While we don't yet know a theory of gravity that describes the black hole in its entirety, we do know how to work with quantum theory on the surface. What that means is that quantum mechanics is valid at a black hole's surface and that as the hole evaporates it doesn't lose information. One caveat is that the type of space-time Maldacena studied is different from the space-time of our universe, but his result is so compelling that physicists are loath to quibble.

if quantum theory breaks down in the vicinity of the event horizon, then it is possible to transfer information from within the black hole to distant outside regions, and thus avoid creating a firewall.

The trouble is, for it to work, Giddings had to relax the law that forbids faster-than-light information transfer. Another problem was that he couldn't say exactly where in space-time quantum theory should break down. Still, it was a tantalising idea.

Enter Polchinski and his students Ahmed Almheiri and James Sully. They reckoned they could crack the problem by combining Giddings's model with earlier work carried out by Leonard Susskind at Stanford University in California.

That meant reconciling the model with three postulates put forward by Susskind, which many physicists hold dear. One is, of course, that information is not lost as a black hole evaporates. The others relate to thought experiments using two observers called Alice and Bob who are approaching a black hole. Intrepid Alice crosses the black hole event horizon. Cautious Bob stays outside.

According to the second postulate, Bob sees nothing unusual as he sits outside the black hole. The third postulate says that Alice also sees nothing amiss as she crosses the event horizon. That's because the event horizon is not a physical boundary, it is just an ordinary patch of vacuum in an ordinary patch of space-time that curves gently.

Polchinski and his colleagues failed in their attempt to reconcile all three – if information wasn't lost, then the firewall still existed and Alice ended up burning to a crisp. But that didn't deter them. "You first try to do something, and if you fail to do it, you try to prove that it is impossible," says Polchinski.

Their colleague Marolf joined them in this new effort and their work led to a paper, published in July 2012, showing that the postulates cannot all be true simultaneously. It caused a storm of controversy; there are already more than 170 papers discussing the work, including one that says the answer is to ignore gravity.

If Hawking radiation does carry quantum information out of the black hole, as many think it does, then quantum mechanics has a few things to say about it. Let's say particle A of Hawking radiation comes out early in the life of the hole. Quantum theory says that particle A shares a spooky connection, or is entangled, with another Hawking radiation particle that emerges later in the life of the black hole.



Hawking radiation causes black holes to evaporate

Now, think about particle B, which exists much later than A. Particle B is one of a pair of particles, B and C, produced at the horizon, and C has fallen into the black hole. Space-time at the horizon is assumed to be nothing special, just gentle gravity and low curvature. This demands that the virtual particles produced at the horizon be entangled with each other. So B must be entangled with C. But since early and late Hawking radiation must be entangled, B is also entangled with A.

Unfortunately this violates another cherished principle of quantum mechanics known as the monogamy of entanglement. Simply speaking, it says that particle B can be entangled with A or C, but not both.

So the conundrum has come full circle. If we want to get information out of a black hole, A must be entangled with B. If we want the vacuum of space-time to be ordinary at the event horizon – which is what allows Alice to slip into the black hole without bursting into flames – then B must be entangled with C.

Something has to give. Is it going to be quantum mechanics or general relativity?

Take quantum mechanics and its prediction that information is conserved. Could this be wrong? Polchinski doesn't think that is possible because of Maldacena's work, which is one of the strongest mathematical statements in favour of leaving quantum mechanics as it is. What's more, quantum mechanics is a theory that has been extremely well tested, and even tiny changes put it at odds with experimental results.

The other option is to call into question the state of the vacuum at the horizon. Messing with monogamy could be avoided if particles B and C on either side of the horizon aren't entangled. But destroying this entanglement leaves the black hole's event horizon in an agitated thermal state and reinstates the firewall. Instead of floating across the event horizon without drama, Alice would face instant incineration by temperatures as high as 10^{32} kelvin.

This has dismayed Marolf. General relativity says crossing a black hole's event horizon should be uneventful. "A firewall would be a strong violation of general relativity," he argues. "In that struggle between general relativity and quantum mechanics, general relativity loses badly. I feel bad about that, because I think of myself as a relativist by training."

Fresh thinking

He's not alone in finding it unpalatable. "You are sailing along fairly fine in this very smooth space-time, and all of a sudden, bam! You hit this firewall and just burn up," says Preskill. "It's pretty crazy."

Nonetheless, the firewall is the best explanation if black holes transfer information into the Hawking radiation. Susskind remains sceptical about the existence of firewalls, but he has argued that they could represent the migration of the singularity, which in traditional black hole physics lies at the centre of the black hole, to the horizon.

Even if firewalls do form, Susskind differs on when they might form. For instance, for a black hole with the radius of a proton, Polchinski, Marolf and colleagues say that firewalls would form 10^{-20} seconds after the formation of the black hole, while in Susskind's view it would take as long as the age of the universe.

Regardless of when firewalls form, if they do, then space-time as we know it may terminate at the horizon. "If the entire black hole horizon becomes this firewall that cuts off the interior, maybe the interior just doesn't exist," says Marolf.

The paradox would also be resolved if there is something special about space-time near a black hole, so that information can be transferred faster than the speed of light. Maybe Giddings and Mathur were on to something, although that would be another blow to relativity.

The upshot of all this is that four decades after Hawking and Bekenstein proposed the black hole information paradox, the problem hasn't gone away. It has just made physicists look deeper at their theories. "I am as confused as I was 20 years ago," says Polchinski.

Preskill says this isn't a bad thing. "There is always a fourth possibility: none of the above, something we haven't thought of," he says. "Any way it shakes down, it's going to be interesting. All the options are crazy, that is what's so wonderful about the situation." ■

"In the struggle between general relativity and quantum mechanics, relativity loses badly"

Firefight at the black hole corral

Trying to tame the cosmic beasts has created a Wild West in physics, finds Lisa Grossman

TAKE what you know about black holes and throw it out. In early 2014, famed physicist Stephen Hawking caused an uproar with his assertion that black holes do not exist – at least not as we’ve defined them for the past 40 years. Rather than letting nothing, not even light, escape their grasp, Hawking says that this “point of no return” is a fallacy, and black holes will sometimes let trapped light back out.

His idea is a proposed solution to the famous black hole firewall paradox (see page 87). This has it that if black holes are as we think, they must be surrounded by rings of fire, although that would violate general relativity.

Hawking’s notion does away with a key part of a black hole – its event horizon. It sounds far out and has been met with fierce debate, but even if Hawking is wrong, his high-profile claim highlights the fact that efforts to solve the paradox have turned cosmology on its ear. Firewalls have turned this corner of physics into a cosmic Wild West: a strange frontier where black holes can take on the form of exotic stars, fuzzballs, time machines or iced wormholes (see right).

The pillars of theoretical physics – quantum mechanics and general relativity – are in a stand-off. One of them will have to blink if this paradox is to be undone.

Trouble began brewing in the 1970s when Hawking mixed quantum mechanics into relativistic black hole theory and concluded

that they should emit a tiny amount of radiation, which steals mass until the black hole evaporates. But that would erase all information regarding the matter that formed the black hole, violating quantum rules.

Although theories did emerge showing how information might escape in a quantum-friendly way, the uneasy peace didn’t last. In 2012, physicists pointed out that an information leak would create scorching-hot particles at the event horizon, which would rapidly incinerate anything falling in. This is anathema to relativity, which says that an observer falling into a black hole shouldn’t notice anything happening. So once again, black holes are up for grabs.

“There are various ideas, and none is completely concrete,” says Juan Maldacena at the Institute for Advanced Study in Princeton, New Jersey. “Hopefully trying to resolve them will lead to a new theory that combines all the ideas into a consistent framework.”

Or we could go back to the original paradox, argues Sabine Hossenfelder at the Nordic Institute for Theoretical Physics in Stockholm, Sweden. Alternative interpretations of quantum theory erase the firewall, but only if information remains trapped.

“No matter which way this goes, it’s clear that the things in conflict are very fundamental,” says Raphael Bousso at the University of California, Berkeley. “You can’t resolve this conflict without a revolution.” ■

“Black holes do not exist, at least not as we have defined them for the past 40 years”

THE FIREWALL PARADOX

Theory says that black holes should have moats of fire around them, but this violates general relativity. Trying to solve the problem has created a herd of wild ideas...

BOUNCING STAR

Instead of evaporating, black holes could end their lives as exotic stars. Carlo Rovelli at the University of Aix-Marseille in France imagined what would happen if a collapsing star bounced back outwards an instant before it would otherwise have become a traditional black hole.

General relativity says that time runs more slowly in the presence of strong gravity, so while the star’s bounce might seem instantaneous to an observer on its surface, from outside we see a black hole that lives for billions of years. Eventually the star pops back out, bringing all its information with it. That kills the firewall, but a major flaw is that, to the star, the bounce happens faster than the speed of light.



STRINGY FUZZBALL

String theory is a popular but as yet unproven framework for trying to unite the quantum and relativistic realms. It has champions in the black hole firewall debate.

If you accept the multiple extra dimensions predicted by string theory, black holes are fuzzballs – tangled balls of strings with no well-defined horizon. In the fuzzball description, says Samir Mathur at Ohio State University in Columbus, space-time itself ends just before what looks like the event horizon.

Weirdly, it still looks like a black hole from the outside. “If you don’t have a horizon, you don’t have radiation, you don’t lose information and we are all done,” he says.

TIME MACHINE

The firewall arises because of the high energy needed for information to escape the event horizon, according to the rules of quantum physics.

Juan Maldacena at the Institute for Advanced Study in Princeton, New Jersey, and his colleagues found a way for information to escape more sedately – by letting time inside

the black hole flow backwards. This means that particles escaping from the event horizon can gather the information that’s flowing back outwards almost instantly and carry it away – no flames needed.

However, there is a problem: an observer falling in could notice the time shift when they entered the black hole, violating relativity.

ICE WALL

Leonard Susskind at Stanford University in California and Juan Maldacena had tried to solve the paradox by turning a black hole into a weird type of wormhole that is quantum mechanically linked to its own escaping particles of radiation. This lets information out with no fire.

But Raphael Bousso at the University of California, Berkeley, says

this solution would raise an equally ugly problem: a wall of “ice”. The quantum link forces particles near the wormhole’s horizon into a particular quantum state, “freezing” them and giving away the horizon’s location.

“It’s like having a signpost at the horizon, which is one thing we’re not allowed,” says Bousso.

CHAPTER SIX

TIME

The now delusion

*Do past, present and future exist only inside our heads? **Michael Slezak** investigates*





IMAGINE standing outside the universe. Not just outside space, but outside time too. From this spectacular vantage point, you gaze down upon the universe. At one end you see its beginning: the big bang. At the other, you see... whatever it is that happens there. Somewhere in the middle is you, a minuscule worm: at one end a baby, the other end a corpse. From this impossible perspective, time does not flow, and there is no “now”. Time is static. Immutable. Frozen.

Fantastical as it seems, for most physicists today the universe is just like that. We might think of time flowing from a real past into a not-yet-real future, but our current theories of space and time teach us that past, present and future are all equally real – and fundamentally indistinguishable. Any sense that our “now” is somehow special, or that time flows past it, is an illusion we create in our heads.

Physics, in fact, has killed time as we know it. The question is: do we need it back?

It was Newton who began to stick the knife into now. His laws of motion, formulated late in the 17th century, were the first to capture time in mathematical equations. Soon it was natural to depict motion on a graph with time on a spatial axis. Once that was done, any special, unique point of “now” started to look as subjective as a “here” on a map of space.

Einstein landed the fatal blow at the turn of the 20th century. According to his special

theory of relativity, there is no way to specify events that everyone can agree happen simultaneously. Two events that are both “now” to you will happen at different times for anyone moving at another speed. Other people will see a different now that might contain elements of yours – but equally might not. “You can define it, but people won’t necessarily agree,” says physicist Sean Carroll of the California Institute of Technology in Pasadena.

The result is a picture known as the block universe: the universe seen from that impossible vantage point outside space and time. You can by all means mark what you think is “now” with a red dot, but there is nothing that distinguishes that place from any other, except that you are there. Past and future are no more physically distinguished than left and right. There are things that are closer to you in time, and things that are further away, just as there are things that are near or far away in space. But the idea that time flows past you is just as absurd as the suggestion that space does.

Contradicting reality

George Ellis, a cosmologist at the University of Cape Town, South Africa, does not buy any of that. The block universe contradicts every single experience we have, he says. “It doesn’t represent the passage of time, and that’s one of the most fundamental features of daily life. So it’s a bad model of reality.” What’s more, accept it and any attempt to understand anything about anything becomes meaningless. “The whole process of doing science depends on time rolling on. We make a hypothesis, test it, accept it or reject it – the process rolls on in time.”

Back in 2006, Ellis began to sketch a different picture. His starting point was not relativity, but quantum physics. A strange fact about quantum physics is that indeterminate future outcomes are seemingly governed by probabilities in the present. Quantum objects exist in “superpositions” of more than one state until such time as we measure them, when they adopt one or other of their possible forms. The most notorious illustration is Schrödinger’s cat: locked in a box with a vial of poison whose seal may or may not be intact, ➤

“Our current theories of space and time teach us that past, present and future are all equally real – and fundamentally indistinguishable”

"We live on an edge of the universe, on a surface we call the present that shimmers into existence one moment at a time"



it is simultaneously dead and alive – until you open the box, when it is most definitely one or the other.

Such quantum oddities are a blow to the block-universe conception of an equally real past, present and future, says Ellis. “Even if you know everything about the state of the universe today, you can’t predict what will be tomorrow. The future can’t be real because it’s not even fixed yet.”

For Ellis, quantum physics supplies an objectively defined present moment: it is the boundary between what our experiments have determined and what remains to be determined. We live on the leading edge of a “growing block” universe, on a surface we call the present that shimmers into existence one

moment at a time as quantum measurements are made.

Ellis thinks it is perfectly possible to define this “now” within relativity, too. Einstein’s general theory of relativity, published a decade after his special theory, is a full picture of space and time, describing how a combined space-time is warped by the presence of matter to produce the force we call gravity. If we gathered enough data and had a big enough computer, we might take account of all the space-time distortions of all the galaxies, black holes and other matter in the universe to calculate a 3D surface on which each point is exactly the same age as the point where we are. “Space-time is defined up to then and not beyond,” says Ellis.

This present is still not “now” as we know it, because not everything on this 3D surface happens simultaneously: as demanded by special relativity, if you and I are moving at different speeds on it, we will still disagree on what is happening now. But that doesn’t necessarily matter. Within relativity, things that are causally related to one another happen in the same order from all perspectives, even if individual observers can’t agree on exactly when they happened. “That’s just psychology,” says Ellis. “It makes you feel happy to think this is simultaneous with that, but it doesn’t mean anything for physics.”

There are still wrinkles in the growing block. Quantum theory reveals that the future is indeterminate, and that aspects of the present

are, too: Schrödinger's poor feline is an example, if we don't bother to check whether it is alive or dead. Ellis's solution, formulated with his colleague Tony Rothman in 2009, is that the present is not a solid surface, but one pitted with indeterminacies that gradually solidify into certainties. These gaps in the present are not something we'd notice: the cat notwithstanding, quantum indeterminacies can only occur for very small things, and over small timescales. "The holes in the present aren't going to be big enough to fall through," says Ellis.

He is still working to fill in gaps in the theory, most recently how the creation of the future cascades down from the cosmological to the quantum scale.

Not everyone is convinced he is on the right track. Huw Price, a philosopher of physics at the University of Cambridge, isn't persuaded by the premise of the present as a dividing line between a real past and an unreal future. Even if the future is indeterminate, he says, it still can be real: you might not be able to determine what's on the other side of a mountain from your current location, but that doesn't mean it doesn't exist.

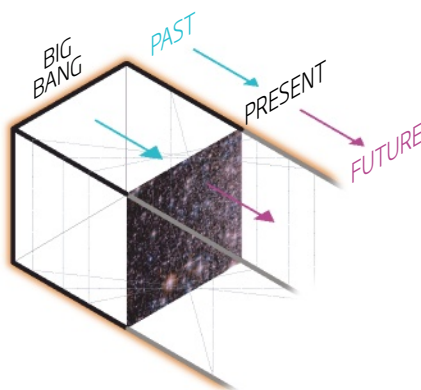
Carroll's beef is that Ellis's argument depends on the truth of the "Copenhagen" interpretation of quantum mechanics. This is the idea that acts of measurement determine the world's future trajectory, and it is the most popular way among physicists to square quantum theory's indeterminacy with the decidedly determined world around us. Carroll prefers the "many worlds" scenario, in which every quantum possibility occurs in different universes: the present Schrödinger's cat is dead in some universes, but alive in others. The cat's future is just as defined as its past; it just has many possible futures. If this interpretation – or any other of the many interpretations of quantum theory – is correct, Ellis's way of defining the present vanishes.

That's one reason why theoretical physicist Lee Smolin of the Perimeter Institute in Waterloo, Ontario, Canada, thinks that we must be more radical to rescue time. In his recent book *Time Reborn*, he argues that if we want to square how we perceive time with what physics tells us about it, it's no good adapting the block universe: we must throw it out altogether.

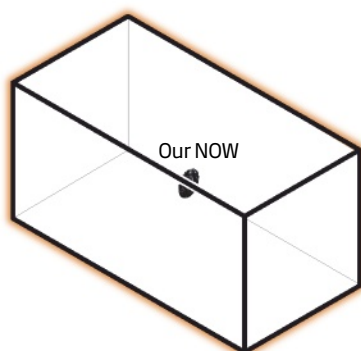
Smolin's starting point is a reformulation of general relativity known as shape dynamics, developed by the independent physicist Julian Barbour and others. Whereas in relativity space and time stretch and condense for observers travelling at different speeds,

About time

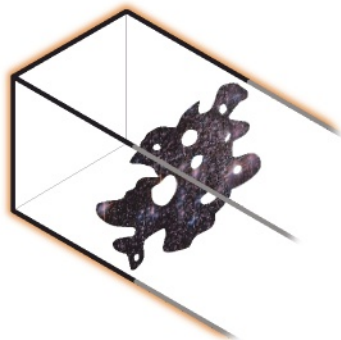
We perceive time as flowing past us from a real past to an as-yet undetermined future. But what's the reality?



Einstein's relativity tells us that time does not flow past a uniquely defined present. Observers elsewhere in the cosmos, or moving at different speeds, will all construct very different nows



There is no possible objective definition of now – all points in space and time form a static "block universe"



In the "growing block" universe, relativity is used to calculate the surface of the universe where everything is the same age. Quantum measurements on this surface continually create the present, one that moves forward in time

in shape dynamics only sizes change. Two distant observers will always agree on what's happening "now" in a galaxy regardless of their relative motions; they just won't be able to agree how big things in that galaxy are.

That might seem like a zero-sum game, replacing one uncomfortable principle with another. For Smolin, though, bending only space, rather than space and time, neatly recreates a conception of time of the sort quantum physics uses, one in which a single external clock provides a beat that distinguishes one moment from the next. The great prize on offer is the possibility of unifying our understanding of quantum theory with that of gravity, the only one of the fundamental forces of nature to have no quantum description. The route to a "theory of everything", Smolin thinks, is through a better understanding of time.

A unique present

Once simultaneity is regained, it becomes possible to describe the entire universe as a series of layered moments – a succession of objectively identified times in which all events are simultaneous. "All that exists is this moment," says Smolin. This is unlike the block universe, where past, present and future are equally real, or Ellis's conception, where only the past and the present are. Instead, the only things that are real about the past or future in Smolin's world are signs of them in the present: records of the past and indicators of what is to come in the future. Smolin is working with Marina Cortes from the University of Edinburgh, UK, to flesh out the idea with mathematics, and to explore which of the many theoretical approaches to quantum gravity it is compatible with.

Price is unmoved. Even if Smolin's or Ellis's approach can provide an objective way of defining the present, he says, there is still a big logical hole. On the one hand, such arguments demand that the present moment be unique; on the other, they demand that every other moment also acquire that unique property. "The whole idea of a privileged present moment is incoherent," he says.

Tim Maudlin, a philosopher and mathematician at New York University, has a different objection. Even if Ellis's or Smolin's theory provides a physical basis for our intuitive conception of here and now, neither explains the fact that we see time flowing, whereas physics suggests it is stationary. This is a fundamental omission, says Maudlin. "The notion that time passes is absolutely ➤

commonplace; it is not a bit of technical jargon invented by philosophers.” Without flowing time, he says, nothing would move at all. Things like rivers appear to flow in space, but “it’s the fundamental direction in time that underlies all of these other directionalities”.

For the past few years, Maudlin has been working on what he calls the theory of linear structures, which he hopes will allow him to reincorporate a flowing time into physics. The idea is rooted in mathematics rather than physics: unlike shape dynamics, it doesn’t provide a rival physical basis for the warped space-time geometry introduced by relativity. “It is the language in which to write a physical theory, not a physical theory itself,” says Maudlin, who has published the idea in book-length detail.

The principal addition to this language’s vocabulary is an object called a directed line. In any conventional geometry, lines between two points in space and time do not come with a natural direction: we have to define a line in terms of a coordinate system, specifying that it passes from me to you rather than you to me, or drawing an arrowhead on the line to make things clear. In Maudlin’s geometrical language, however, that arrowhead is implicit in the definition of any line. Once this is built into the fundamental language of geometry, time can naturally acquire a direction.

Ellis thinks Maudlin’s work is interesting, and also compatible with his growing block

“People who believe in physics know that the distinction between past, present and future is only a stubborn illusion”

picture, explaining in more detail how the flow of time can be fundamental to physics. “In the end, you have to base your theories on some fundamental givens. Time, it’s just kind of a given, which everything else flows around,” he says.

Carroll is more sceptical. Rather than attempting to change the block universe to explain our experience of time flowing, he says we should concentrate on explaining human experience in light of what our very successful physics tells us about the block universe. That task, he says, is quite achievable. “That doesn’t mean that we’ve done it yet, but I see no obstacle to doing it.”

Going with the flow

Craig Callender, a philosopher from the University of California, San Diego, agrees. Explaining our apparently aberrant perception of time does not mean we have to overturn physics or invent a whole new way of doing geometry. When, he says, we “embed critters like us” in a universe like ours, it makes sense that we should see a flowing time and

distinguish past, present and future – even when the reality is something different.

To explain why, we can return to that vantage point gazing down on the entire block universe, and zoom in on that tiny human speck: the four-dimensional worm with a baby at one end and a corpse at the other. This worm’s perception of time differs from “reality” first in that it remembers the past but does not see the future. That can be explained as a result of thermodynamics. The universe started off in a highly ordered, taut state after the big bang, and has been expanding into an ever more disordered, flaccid state ever since. There is an infinitude of paths in which the universe can evolve forward in time, but only one path back into its history. Why the universe works like that is another, fundamentally unanswered question – but it means that, purely statistically, we are only ever likely to have a clear view backwards in time.

Even then, you would expect us worms to feel as if we are stationary in time with a view in only one direction, rather what we experience: moving backwards into the future with no clear view of where we are heading. For Callender, the key to this illusion is an important psychological fact about ourselves: we have a sense of identity. According to physics, your life is described by a series of slices of your worm – you as a baby, you as you ate breakfast this morning, you as you started reading this sentence and so on, with each slice existing motionless in its respective time. We generate time’s flow by thinking that the same self that ate breakfast this morning also started reading this sentence. “Really there’s all these different mes at all these different times,” says Callender. “But because I think that I’m identical over time, that’s why time seems to flow, even though it doesn’t.”

So do we really need to mourn time’s passing? Einstein, for one, drew solace from the view of the timeless universe he had helped to create, consoling the family of a recently deceased friend: “Now he has departed from this strange world a little ahead of me. That means nothing. People like us, who believe in physics, know that the distinction between past, present and future is only a stubbornly persistent illusion.” So it goes. ■



Is time in need of repair, or can the universe cope without it?

Are we nearly there yet?

The quest for an ultimate theory might have stalled – but let's not call the journey off, says **Michael Brooks**

BLAME the ancient Greeks: they started it. The idea seemed sensible enough. If we dig down, we can uncover the building blocks of reality: what matter is ultimately made of, what rules govern its behaviour. Dig down far enough and we will hit the ultimate gold – a theory of how everything works.

In some senses, we have been doing quite well. Quantum theory's weirdness might bamboozle us, but the standard model of particle physics built upon it successfully reduces things to a few fundamental particles and just three forces. General relativity's mind-warping treatment of space and time delivers a

cosmology that describes with stunning accuracy a universe dominated by a fourth force, gravity. Yes, there are wrinkles in both theories, but surely they will be ironed out as we move towards an ultimate theory of physics that unifies the two?

Put that to many physicists, and you will get a grumpy response. "Elementary particle physics is pretty much in the state chemistry was in with Dmitri Mendeleev," during the creation of the periodic table, says David Deutsch at the University of Oxford. "It has been classifying things, and has recognised there is an underlying structure, but it has no idea what that underlying structure is."

Crucially, quantum theory and general relativity remain fundamentally incompatible. That is not a problem when we use relativity to describe the very large – stars, galaxies, the cosmos – and quantum theory for the very small – molecules, atoms, subatomic particles. But for a full understanding of the universe we have to know how the tiny newborn universe became so big: going back towards the big bang requires both theories to work together. So too, possibly, does the existence of black holes. As Stephen Hawking and Jacob Bekenstein showed in the 1970s, these general-relativistic monsters might ➤

destroy information, something forbidden by quantum theory.

Even something as basic as space and time highlights how badly these two theories get along. Relativity's space-time is a smooth four-dimensional sheet; the quantum field theories that underlie the standard model suggest space is pixelated into units with sizes of about 10^{-35} metres, and do not even treat time as a real and observable thing.

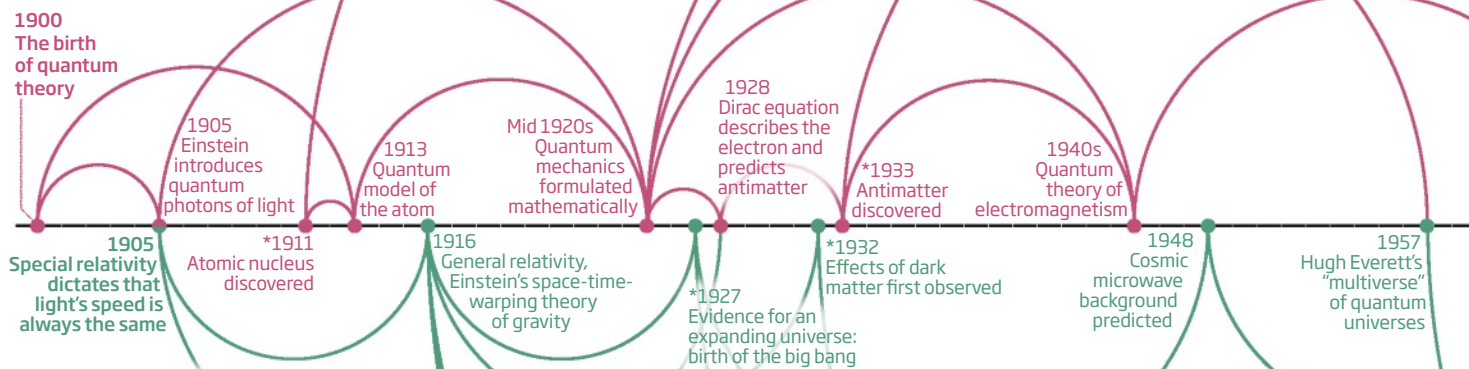
Asked to choose between the two

candidate for an ultimate theory. String theory predicts that space has hidden extra dimensions, invoking symmetries embedded in these dimensions to "fold" energy into geometric shapes that look like certain fundamental particles, or mimic the way space curves in the presence of mass.

The theory has produced some credible depictions of particles, among them the long-sought graviton, a quantum particle that would carry the force of gravity. It

compelling", he says, but there is no reason to think one exists – or that we can find it. That we have got so far with mathematics is a remarkable fact, but it does not mean we can go all the way.

One problem is that mathematics provides infinite ways in which numbers and abstract quantities can be processed – but no indication of what exists beyond it. "Mathematics only reveals truth about abstract objects," says Deutsch. "Physics is not so much trying to study those



Our standard models of particle physics and cosmology have been built on a series of interlinked theoretical and experimental discoveries (*)

theories, most physicists' money is on quantum theory being "right", because its mathematics is such a successful prism through which to view the world. Others, from Einstein onwards, have taken issue with quantum theory's seeming "irreality" and spooky, counter-intuitive correlations between apparently unrelated objects. If we cannot find a convincing physical reason why these correlations are just so, they argue, perhaps quantum theory is just an approximation to something better.

Attempts to go beyond this impasse have drawn on favoured mathematical ideas such as symmetry. One result is supersymmetry, a theory widely regarded as a way station on the road to string theory, a favoured

thus takes steps towards a unified picture of all four forces of nature on the basis of quantum theory. But like other proposed frameworks for a theory of everything (see "Six routes to an ultimate theory", right), it has a big flaw. "String theory does predict new things, but they're almost certainly not testable in the foreseeable future," says Paul Davies of Arizona State University in Tempe.

That failing means the idea of a theory of everything has quietly disappeared, says Renate Loll of Radboud University in Nijmegen in the Netherlands. "For a while you would see it in papers, in the heyday of string theory, but it has gone totally out of fashion." Chris Isham of Imperial College London goes further. A theory of everything is "psychologically

objects as to find out which one of them corresponds to reality." The fraction of the pure mathematics we know about that we have used so far to construct physical theories is relatively small.

For example, all the relationships between particles and fields and space and time can be represented by a subset of mathematical operations that are computable on a Turing machine, the basis of all our computers. "We don't know why," says Deutsch. "It just seems to be a brute fact." Progress towards a theory of everything may require branching into areas of mathematics that today are not regarded as computable.

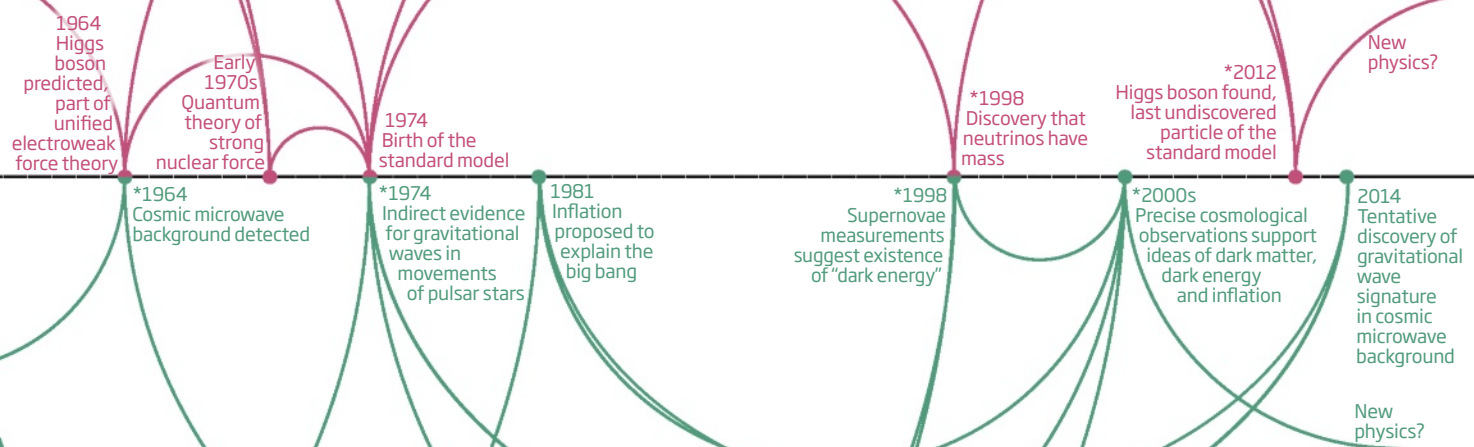
Deutsch's own hunch is even more radical. He says we must abandon the idea that has served us so well over the past

century: that if we start with the mathematics, reality will follow. Instead we must look first to problems in our understanding of the physical universe – the universe’s missing matter, say, or why gravity is so much weaker than the other forces – and try to work out what changes to our world view might solve it. “Then you have to formulate it mathematically, and test it. It has to be in that order,” says Deutsch. “Quite a lot of theoretical physics tries to do the mathematics first,

and that’s never going to work.”

Is it worth the effort? Davies thinks so – but only if we are clear that even a theory of everything, at least as physicists define it, will not contain all the answers. “It wouldn’t help solve problems of the origin of life or the nature of consciousness,” he says. Lisa Randall of Harvard University puts things even more bluntly: “Even if we knew the ultimate underlying theory, how are you going to explain the fact that we’re here?”

Perhaps the greatest reason to continue the search, though, lies not in the goal but in the routes we take to reach it. Our greatest, life-changing scientific insights have come from attempts to simplify physics and bring together disparate areas of understanding. James Clerk Maxwell united our understanding of electricity and magnetism, providing the theoretical bedrock on which most modern technology has been built. Einstein unified ideas about energy and



SIX ROUTES TO AN ULTIMATE THEORY

STRING THEORY

Particles and forces exist as tiny vibrating strings. Different string vibrations produce particles with different characteristics: electrons, quarks, the Higgs boson. The action takes place in 10 (mainly tiny) dimensions.

M-THEORY

A more general 11-dimensional theory that ties together variants of string theory.

LOOP QUANTUM GRAVITY

Breaks space up into small chunks

and uses knot theory to tangle up links between them, giving these links the characteristics of elementary particles.

CAUSAL DYNAMICAL TRIANGULATION

Proceeds in a similar way to loop quantum gravity, but involves time as well as space.

AN EXCEPTIONALLY SIMPLE THEORY OF EVERYTHING

A 2007 paper by the independent researcher Garrett Lisi caused a rumpus by claiming a unified

theory of all forces lies in the symmetries of a mathematical object known as the E8 Lie group. But even its author admits it is so far little more than a “speculative proposal”.

INFORMATION THEORY

Thermodynamics rules everything, including the flow of information. A better understanding of how might provide clues as to why things are as they are – and narrow down the parameters of any ultimate theory.

mass to come up with $E=mc^2$ – opening up the nuclear age, among other things. “Historically, these things have usually led somewhere,” says Davies.

For a purist, this might be an excessively utilitarian take. But anyone who searches for an ultimate theory purely for its own sake risks ultimate disappointment, says Davies, and also risks making the same mistake as those who believed at the end of the 19th century that physics was complete. “You could come up with some marvellous scheme, and you could put it up in a stained-glass window and celebrate it as a wonderful achievement of the human intellect. But there would always be the possibility that someone comes up with a better one.” ■



SIMON DANAWER

IN FROM THE COLD



Forget quantum physics, forget relativity. Inkling of an ultimate theory might emerge from an unexpected place, says physicist [Vlatko Vedral](#)

AS REVOLUTIONS go, its origins were haphazard. It was, according to the ringleader Max Planck, an “act of desperation”. In 1900, he proposed the idea that energy comes in discrete chunks, or quanta, simply because the smooth delineations of classical physics could not explain the spectrum of energy re-radiated by an absorbing body.

Yet rarely was a revolution so absolute. Within a decade or so, the cast-iron laws that had underpinned physics since Newton’s day were swept away. Classical certainty ceded its stewardship of reality to the probabilistic rule of quantum mechanics, even as the parallel revolution of Einstein’s relativity displaced our cherished, absolute notions of space and time. This was complete regime change.

Except for one thing. A single relic of the old order remained, one that neither Planck nor Einstein nor any of their contemporaries had the will or means to remove. The British astrophysicist Arthur Eddington summed up the situation in 1915. “If your theory is found to be against the second law of thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation,” he wrote.

In this essay, I will explore the fascinating question of why the laws of thermodynamics have proved so formidably robust since their origins in the early 19th century. The journey traces the deep connections that were discovered in the 20th century between thermodynamics and information theory – connections that allow us to trace intimate links between thermodynamics and not only quantum theory but also, more speculatively, relativity. Ultimately, I will argue, those links show us how thermodynamics in the 21st century can guide us towards a theory that will supersede them both.

In its origins, thermodynamics is a theory about heat: how it flows and what it can be made to do. The French engineer Sadi Carnot formulated the second law in 1824 to

characterise the mundane fact that the steam engines then powering the industrial revolution could never be perfectly efficient. Some of the heat you pumped into them always flowed into the cooler environment, rather than staying in the engine to do useful work. That is an expression of a more general rule: unless you do something to stop it, heat will naturally flow from hotter places to cooler places to even up any temperature differences it finds. The same principle explains why keeping the refrigerator in your kitchen cold means pumping energy into it; only that will keep warmth from the surroundings at bay.

A few decades after Carnot, the German physicist Rudolph Clausius explained such phenomena in terms of a quantity that he called entropy, characterising disorder. In this picture, the universe works on the back of processes that increase entropy – for example dissipating heat from places where it is concentrated, and therefore more ordered, to cooler areas, where it is not.

That predicts a grim fate for the universe itself. Once all heat is maximally dissipated, no useful process can happen in it any more: it dies a “heat death”. A perplexing question is raised at the other end of cosmic history, too. If nature always favours states of high entropy, how and why did the universe start in a state that seems to have been of comparatively low entropy? At present we have no answer, and later I will mention an intriguing alternative view.

Perhaps because of such undesirable

consequences, the legitimacy of the second law was for a long time questioned. The charge was formulated with the most striking clarity by the British physicist James Clerk Maxwell in 1867. He was satisfied that inanimate matter presented no difficulty for the second law. In an isolated system, heat always passes from the hotter to the cooler, and a neat clump of dye molecules readily dissolves in water and disperses randomly, never the other way round. Disorder as embodied by entropy does always increase.

Maxwell’s problem was with life. Living things have “intentionality”: they deliberately do things to other things to make life easier for themselves. Conceivably, they might try to reduce the entropy of their surroundings and thereby violate the second law.

Information is power

Such a possibility is highly disturbing to physicists: something is either a universal law or it is merely a cover for something deeper. Yet it was only in the late 1970s that Maxwell’s entropy-fiddling “demon” was laid to rest. Its slayer was the US physicist Charles Bennett, who built on work by his colleague at IBM, Rolf Landauer, using the theory of information developed a few decades earlier by Claude Shannon. An intelligent being can certainly rearrange things to lower the entropy of its environment. But to do this, it must first fill up its memory, gaining information as to how things are arranged in the first place.

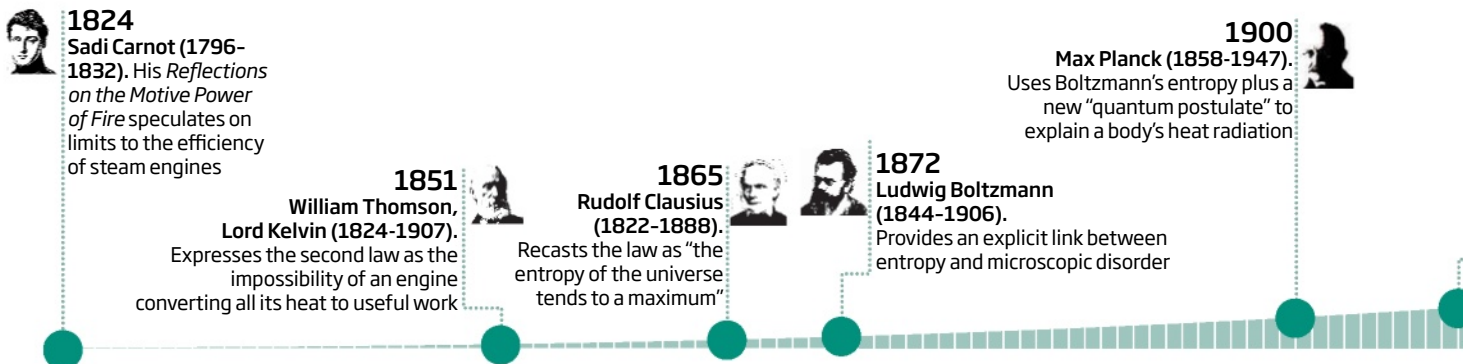
This acquired information must be encoded somewhere, presumably in the demon’s memory. When this memory is finally full, or the being dies or otherwise expires, it must be reset. Dumping all this stored, ordered information back into the environment increases entropy – and this entropy increase, Bennett showed, will ultimately always be at least as large as the entropy reduction the demon originally achieved. Thus the status of the second law was assured, albeit anchored in a mantra of Landauer’s that would have been unintelligible to the 19th-century progenitors of thermodynamics: that “information is physical”.

But how does this explain that thermodynamics survived the quantum revolution? Classical objects behave very differently to quantum ones, so the same is presumably true of classical and quantum information. After all, quantum computers are notoriously more powerful than classical ones (or would be if realised on a large scale).

The reason is subtle, and it lies in a connection between entropy and probability contained in perhaps the most profound and beautiful formula in all of science. Engraved on the tomb of the Austrian physicist Ludwig Boltzmann in Vienna’s central cemetery, it reads simply $S = k \log W$. Here S is entropy – the macroscopic, measurable entropy of a gas, for example – while k is a constant of nature that today bears Boltzmann’s name. $\log W$ is the mathematical logarithm of a microscopic, probabilistic quantity W – in a gas, this would

Entropy increasing

From a tightly defined origin in the workings of steam engines, the influence of the second law of thermodynamics has steadily spread over the past 200 years





Penetrating insight:
Ludwig Boltzmann's
tomb in Vienna

be the number of ways the positions and velocities of its many individual atoms can be arranged.

On a philosophical level, Boltzmann's formula embodies the spirit of reductionism: the idea that we can, at least in principle, reduce our outward knowledge of a system's activities to basic, microscopic physical laws. On a practical, physical level, it tells us that all we need to understand disorder and its increase is probabilities. Tot up the number of configurations the atoms of a system can be in and work out their probabilities, and what emerges is nothing other than the entropy that determines its thermodynamical behaviour. The equation asks no further questions about the nature of the underlying laws; we need not care if the dynamical processes that create the probabilities are classical or quantum in origin.

There is an important additional point to be made here. Probabilities are fundamentally different things in classical and quantum physics. In classical physics they are "subjective" quantities that constantly change as our state of knowledge changes. The probability that a coin toss will result in heads or tails, for instance, jumps from $\frac{1}{2}$ to 1 when we observe the outcome. If there were a being who knew all the positions and momenta of all the particles in the universe – known as a "Laplace demon", after the French mathematician Pierre-Simon Laplace, who first countenanced the possibility – it would be able to determine the course of all

"Something is either
a fundamental law
or is merely a cover
for something deeper"



1909
Constantin Carathéodory
(1873-1950). Rewrites the
second law in logical terms of
allowed and forbidden processes



1948
Claude Shannon (1916-2001).
Introduces "information entropy" as
a measure of uncertainty in an
encoded message



1961
Rolf Landauer (1927-1999).
Shows that any information-
destroying process also
increases physical entropy



1982
Charles Bennett (1943-).
Uses Landauer's principle to
show an intelligent being
cannot buck the second law



1995
Ted Jacobson (1954-).
Suggests that gravity
is merely an effect of
bodies increasing
their entropy

subsequent events in a classical universe, and would have no need for probabilities to describe them.

In quantum physics, however, probabilities arise from a genuine uncertainty about how the world works. States of physical systems in quantum theory are represented in what the quantum pioneer Erwin Schrödinger called catalogues of information, but they are catalogues in which adding information on one page blurs or scrubs it out on another. Knowing the position of a particle more precisely means knowing less well how it is moving, for example. Quantum probabilities are “objective”, in the sense that they cannot be entirely removed by gaining more information.

That casts in an intriguing light thermodynamics as originally, classically, formulated. There, the second law is little

“Not only should future physics conform to thermodynamics, it should be constructed in its image”

more than impotence written down in the form of an equation. It has no deep physical origin itself, but is an empirical bolt-on to express the otherwise unaccountable fact that we cannot know, predict or bring about everything that might happen, as classical dynamical laws suggest we can. But this changes as soon as you bring quantum physics into the picture, with its attendant notion that uncertainty is seemingly hardwired into the fabric of reality. Rooted in probabilities, entropy and thermodynamics acquire a new, more fundamental physical anchor.

It is worth pointing out, too, that this deep-rooted connection seems to be much more general. Together with my colleagues Markus Müller of the University of Heidelberg, Germany, and Oscar Dahlsten at the University of Oxford, I have looked at what happens to thermodynamical relations in a generalised class of probabilistic theories that embrace quantum theory and much more besides. There too, the crucial relationship between information and disorder, as quantified by entropy, survives.

One theory to rule them all

As for gravity – the only one of nature’s four fundamental forces not covered by quantum theory – a more speculative body of research suggests it might be little more than entropy in disguise (see “Falling into disorder”, left). If so, that would also bring Einstein’s general theory of relativity, with which we currently describe gravity, firmly within the purview of thermodynamics.

Take all this together, and we begin to have a hint of what makes thermodynamics so successful. The principles of thermodynamics are at their roots all to do with information theory. Information theory is simply an embodiment of how we interact with the universe – among other things, to construct theories to further our understanding of it. Thermodynamics is, in Einstein’s term, a “meta-theory”: one constructed from principles over and above the structure of any dynamical laws we devise to describe reality’s workings. In that sense we can argue that it is more fundamental than either quantum physics or general relativity.

If we can accept this and, like Eddington and his ilk, put all our trust in the laws of thermodynamics, I believe it may even afford us a glimpse beyond the current physical order. It seems unlikely that quantum physics and relativity represent the last revolutions in physics. New evidence could at any time foment their overthrow. Thermodynamics

FALLING INTO DISORDER

Although thermodynamics seems to float above the precise content of the physical world it describes, whether classical, quantum or post-quantum (see main story), its connection with the other pillar of modern physics, general relativity, might be more direct. General relativity describes the force of gravity. In 1995, Ted Jacobson of the University of Maryland in College Park claimed that gravity could be a consequence of disorder as quantified by entropy.

His mathematical argument is surprisingly simple, but rests on two disputed theoretical relationships. The first was argued in the early 1970s by Jacob Bekenstein, who was examining the fate of the information in a body gulped by a black hole. This is a naked challenge to the universal validity of thermodynamics: any increase in disorder in the cosmos could be reversed by throwing the affected

system into a black hole.

Bekenstein showed that this would be countered if the black hole simply grew in area in proportion to the entropy of the body it was swallowing. Then each tiny part of its surface would correspond to one bit of information that still counts in the universe’s ledger. This relationship has since been elevated to the status of a principle, the holographic principle, that is supported by a host of other theoretical ideas – but not as yet by any experiment.

The second relationship is a suggestion by Paul Davies and William Unruh, also first made in the 1970s, that an accelerating body radiates tiny amounts of heat. A thermometer waved around in a perfect vacuum, where there are no moving atoms that can provide us with a normal conception of temperature, will record a non-zero temperature. This is an attractive yet counter-intuitive

idea, but accelerations far beyond what can presently be achieved are required to generate enough radiation to test it experimentally.

Put these two speculative relations together with standard, undisputed connections between entropy, temperature, kinetic energy and velocity, and it is possible to construct a quantity that mathematically looks like gravity, but is defined in terms of entropy.

Others, such as Erik Verlinde of the University of Amsterdam in the Netherlands, have since been tempted down the same route.

Such theories, which are by no means universally accepted, suggest that when bodies fall together it is not the effect of a separate fundamental force called gravity, but because the heating that results best fulfils the thermodynamic dictat that entropy in the universe must always increase.



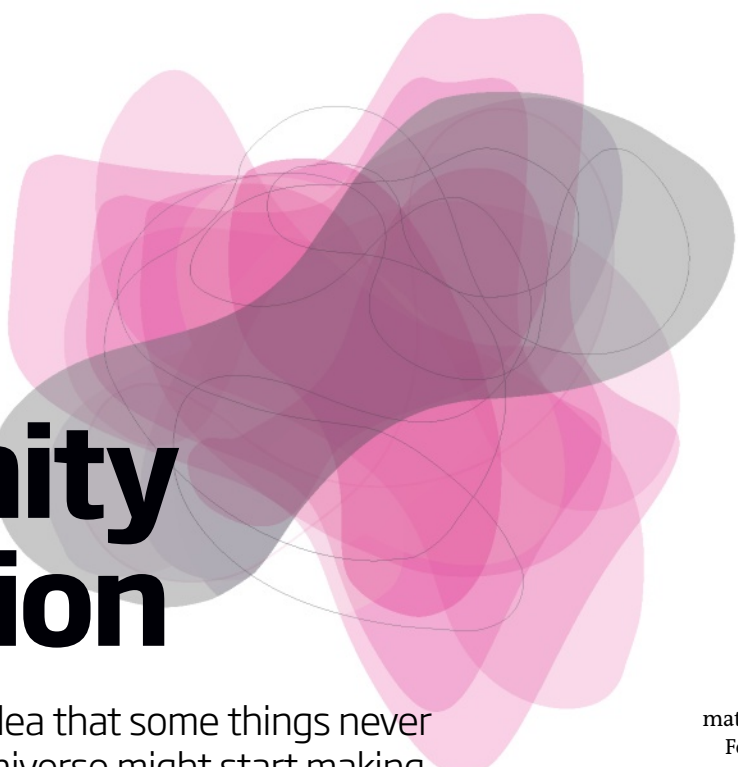
SIMON DANAHER

My colleague at the University of Oxford, the physicist David Deutsch, thinks we should take things much further. Not only should any future physics conform to thermodynamics, but the whole of physics should be constructed in its image. The idea is to generalise the logic of the second law as it was stringently formulated by the mathematician Constantin Carathéodory in 1909: that in the

James Joule's 19th-century experiments with beer can be used to illustrate this idea. The English brewer, whose name lives on in the standard unit of energy, sealed beer in a thermally isolated tub containing a paddle wheel that was connected to weights falling under gravity outside. The wheel's rotation warmed the beer, increasing the disorder of its molecules and therefore its entropy. But hard as we might try, we simply cannot use Joule's set-up to decrease the beer's temperature, even by a fraction of a millikelvin. Cooler beer is, in this instance, a state regrettably beyond the reach of physics.

The question is whether we can express the whole of physics simply by enumerating possible and impossible processes in a given situation. This is very different from how physics is usually phrased, in both the classical

Such an approach would probably please Einstein, who once said: “What really interests me is whether God had any choice in the creation of the world.” A thermodynamically inspired formulation of physics might not answer that question directly, but leaves God with no choice but to be a thermodynamicist. That would be a singular accolade for those 19th-century masters of steam: that they stumbled upon the essence of the universe, entirely by accident. The triumph of thermodynamics would then be a revolution by stealth, 200 years in the making. ■



The infinity illusion

Abandon the idea that some things never end, and the universe might start making more sense, says **Amanda Gefter**

INFINITY. It is a concept that defies imagination. We have a hard-enough time trying to wrap our minds around things that are merely extremely big: our solar system, our galaxy, the observable universe. But those scales are nothing compared with the infinite. Just thinking about it can make you queasy.

But we cannot avoid it. Mathematics as we know it is riddled with infinities. The number line stretches to eternity and beyond, and is infinitely divisible: countless more numbers lurk between any two others. The number of digits in a constant like pi is limitless. Whether geometry, trigonometry or calculus, the mathematical manipulations we use to make sense of the world are built on the idea that some things never end.

Trouble is, once unleashed these infinities are wild, unruly beasts. They blow up the equations with which physicists attempt to explain nature's fundamentals. They obstruct a unified view of the forces that shape the cosmos. Worst of all, add infinities to the explosive mixture that made up the infant universe and they prevent us from making any scientific predictions at all.

All of which encourages a bold speculation among a few physicists and mathematicians: can we do away with infinity?

Belief in the never-ending has not always been a mainstream view. "For most of the

history of mathematics, infinity was kept at arm's length," says mathematician Norman Wildberger of the University of New South Wales in Sydney, Australia. For greats of the subject, from Aristotle to Newton and Gauss, the only infinity was a "potential" infinity. This type of infinity allows us to add 1 to any number without fear of hitting the end of the number line, but is never actually reached itself. That is a long way from accepting "actual" infinity – one that has already been reached and conveniently packaged as a mathematical entity we can manipulate in equations.

Things changed in the late 19th century, when the German mathematician Georg Cantor invented set theory, the underpinning of modern number theory. He argued that sets containing an infinite number of elements were themselves mathematical objects. This masterstroke allowed the meaning of numbers to be pinned down in a rigorous way that had long eluded mathematicians. Within set theory, the infinite continuum of the "real" numbers, including all the rational numbers (those, like $\frac{1}{2}$, which can be expressed as a ratio of integers) and the irrational numbers (those that cannot, like pi) came to be treated as actual, rather than potential, infinities. "No one shall expel us from the paradise Cantor has created," the

mathematician David Hilbert later declared.

For physicists, however, the infinite paradise has become more like purgatory. To take one example, the standard model of particle physics was long beset by pathological infinities, for instance in quantum electrodynamics, the quantum theory of the electromagnetic force. It initially showed the mass and charge of an electron to be infinite.

Decades of work, rewarded by many a Nobel prize, banished these nonsensical infinities – or most of them. Gravity has notoriously resisted unification with the other forces of nature within the standard model, seemingly immune to physicists' best tricks for neutralising infinity's effects. In extreme circumstances such as in a black hole's belly, Einstein's equations of general relativity, which describe gravity's workings, break down as matter becomes infinitely dense and hot, and space-time infinitely warped.

But it is at the big bang that infinity wreaks the most havoc. According to the theory of cosmic inflation, the universe underwent a burst of rapid expansion in its first fraction of a second. Inflation explains essential features of the universe, including the existence of stars and galaxies. But it cannot be stopped. It continues inflating other bits of space-time long after our universe has settled down, creating an infinite "multiverse" in an eternal stream of big bangs. In an infinite multiverse, everything that can happen will happen an infinite number of times. Such a cosmology predicts everything – which is to say, nothing.

This disaster is known as the measure problem, because most cosmologists believe it will be fixed with the right "probability measure" that would tell us how likely we are to end up in a particular sort of universe and ➤



so restore our predictive powers. Others think there is something more fundamental amiss. "Inflation is saying, hey, there's something totally screwed up with what we're doing," says cosmologist Max Tegmark of the Massachusetts Institute of Technology. "There's something very basic we've assumed that's just wrong."

For Tegmark, that something is infinity. Physicists treat space-time as an infinitely stretchable mathematical continuum; like the line of real numbers, it has no gaps. Abandon that assumption and the whole cosmic story changes. Inflation will stretch space-time only until it snaps. Inflation is then forced to end, leaving a large, but finite, multiverse. "All of our problems with inflation and the measure problem come immediately from our assumption of the infinite," says Tegmark. "It's the ultimate untested assumption."

Disruptive influence

There are also good reasons to think it is an unwarranted one. Studies of the quantum properties of black holes by Stephen Hawking and Jacob Bekenstein in the 1970s led to the development of the holographic principle, which makes the maximum amount of information that can fit into any volume of space-time proportional to roughly one-quarter the area of its horizon. The largest number of informational bits a universe of our size can hold is about 10^{122} . If the universe is indeed governed by the holographic principle, there is simply not enough room for infinity.

Certainly we need nothing like that number of bits to record the outcome of experiments. David Wineland, a physicist at the National Institute of Standards and Technology in Boulder, Colorado, shared the 2012 Nobel prize in physics for the world's most accurate measuring device, an atomic clock that can measure increments of time out to 17 decimal places. The electron's anomalous magnetic moment, a measure of tiny quantum effects on the particle's spin, has been measured out to 14 decimal places. But even the best device will never measure with infinite accuracy, and that makes some physicists very itchy. "I don't think anyone likes infinity," says Raphael Bousso of the University of California at Berkeley. "It's not the outcome of any experiment."

But if infinity is such an essential part of mathematics, the language we use to describe the world, how can we hope to get rid of it?

Sometimes what seems to be infinite is simply very, very long

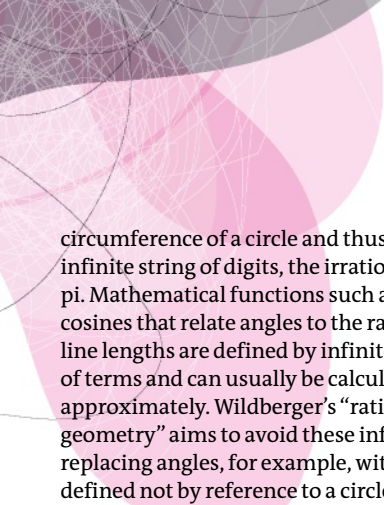
Wildberger has been trying to figure that out, spurred on by what he sees as infinity's disruptive influence on his own subject. "Modern mathematics has some serious logical weaknesses that are associated in one way or another with infinite sets or real numbers," he says.

For the past decade, he has been working on a new, infinity-free version of trigonometry and Euclidean geometry. In standard trigonometry, the infinite is ever-present. Angles are defined by reference to the

"Inflation is saying, hey, there's something totally screwed up with what we're doing. Some basic assumption is very wrong"



CEDRIC LEFEBVRE/GALLERY STOCK



circumference of a circle and thus to an infinite string of digits, the irrational number pi. Mathematical functions such as sines and cosines that relate angles to the ratios of two line lengths are defined by infinite numbers of terms and can usually be calculated only approximately. Wildberger's "rational geometry" aims to avoid these infinities, replacing angles, for example, with a "spread" defined not by reference to a circle, but as a rational output extracted from mathematical vectors representing two lines in space.

Doron Zeilberger of Rutgers University in Piscataway, New Jersey, thinks the work has potential. "Everything is made completely rational. It's a beautiful approach," he says.

Then again, Zeilberger himself subscribes to a view of infinity so radical that it would have even the pre-Cantor greats of mathematics stirring in their coffins. While Wildberger's work is concerned with doing away with actual infinity as a real object used in mathematical manipulations, Zeilberger wants to dispose of potential infinity as well. Forget everything you thought you knew about mathematics: there is a largest number. Start at 1 and just keep on counting and eventually you will hit a number you cannot exceed – a kind of speed of light for mathematics.

That raises a host of questions. How big is the biggest number? "It's so big you could never reach it," says Zeilberger. "We don't know what it is so we have to give it a name, a symbol. I call it N_0 ." What happens if you add 1 to it? Zeilberger's answer comes by analogy to a computer processor. Every computer has a largest integer number that it can handle: exceed it, and you will either get an "overflow error" or the processor will reset the number to zero. Zeilberger finds the second option more elegant. Enough of the number line, stretching infinitely far in both directions. "We can redo mathematics postulating that there is a biggest number and make it circular," he says.

Hugh Woodin, a set theorist at the University of California, Berkeley, is sceptical. "He could be correct, of course. But to me the view is a limiting view. Why take it unless one has strong evidence that it is correct?" For him, the success of set theory with all its infinities is reason enough to defend the status quo.

So far, finitist mathematics has received most attention from computer scientists and robotics researchers, who work with finite forms of mathematics as a matter of course. Finite computer processors cannot actually deal with real numbers in their full infinite glory. They approximate them using floating-point arithmetic – a form of scientific

notation that allows the computer to drop digits from a real number, and so save on memory without losing its overall scope.

The idea that our finite universe might work similarly has a history. Konrad Zuse, a German engineer and one of the pioneers of floating-point arithmetic, built the world's first programmable electronic computer in his parents' living room in 1938. Seeing that his machine could solve differential equations (which ordinarily use infinitely small steps to calculate the evolution of a physical system) without recourse to the infinite, he was persuaded that continuous mathematics was just an approximation of a discrete and finite reality. In 1969, Zuse wrote a book called *Calculating Space* in which he argued that the universe itself is a digital computer – one with no room for infinity.

Tegmark, for his part, is intrigued by the fact that the calculations and simulations that physicists use to check a theory against the hard facts of the world can all be done on a finite computer. "That already shows that we don't need the infinite for anything we're doing," he says. "There's absolutely no evidence whatsoever that nature is doing it

"Start at 1 and just keep on counting and eventually you will hit a number you cannot exceed – a speed of light for mathematics"

any differently, that nature needs to process an infinite amount of information."

Seth Lloyd, a physicist and quantum information expert also at MIT, counsels caution with such analogies between the cosmos and an ordinary, finite computer. "We have no evidence that the universe behaves as if it were a classical computer," he says, "and plenty of evidence that it behaves like a quantum computer."

At first glance, that would seem to be no problem for those wishing to banish infinity. Quantum physics was born when, at the turn of the 20th century, physicist Max Planck showed how to deal with another nonsensical infinity. Classical theories were indicating that the amount of energy emitted by a perfectly absorbing and radiating body should be infinite, which clearly was not the case. Planck solved the problem by suggesting that energy comes not as an infinitely divisible continuum, but in discrete chunks – quanta.

The difficulties start with Schrödinger's cat. When no one is watching, the famous quantum feline can be both dead and alive at the same time: it hovers in a "superposition" of multiple, mutually exclusive states that blend together continuously. Mathematically, this continuum can only be depicted using infinities. The same is true of a quantum computer's "qubits", which can perform vast numbers of mutually exclusive calculations simultaneously, just as long as no one is demanding an output. "If you really wanted to specify the full state of one qubit, it would require an infinite amount of information," says Lloyd.

Down the rabbit hole

Tegmark is unfazed. "When quantum mechanics was discovered, we realised that classical mechanics was just an approximation," he says. "I think another revolution is going to take place, and we'll see that continuous quantum mechanics is itself just an approximation to some deeper theory, which is totally finite."

Lloyd counters that we ought to work with what we have. "My feeling is, why don't we just accept what quantum mechanics is telling us, rather than imposing our prejudices on the universe? That never works," he says.

For physicists looking for a way forward, however, it is easy to see the appeal. If only we could banish infinity from the underlying mathematics, perhaps we might see the way to unify physics. For Tegmark's particular bugbear, the measure problem, we would be freed from the need to find an arbitrary probability measure to restore cosmology's predictive power. In a finite multiverse, we could just count the possibilities. If there really were a largest number, then we would only have to count so high.

Woodin would rather separate the two issues of physical and mathematical infinities. "It may well be that physics is completely finite," he says. "But in that case, our conception of set theory represents the discovery of a truth that is somehow far beyond the physical universe."

Tegmark, on the other hand, thinks the mathematical and physical are inextricably linked – the further we plunge down the rabbit hole of physics to deeper levels of reality, the more things seem to be made purely of mathematics. For him, the fatal error message contained in the measure problem is saying that if we want to rid the physical universe of infinity, we must reboot mathematics, too. "It's telling us that things aren't just a little wrong, but terribly wrong." ■



Could ethereal neutrinos guide us towards the new physics we are so desperately seeking?
Robert Adler investigates

Ghost matter

UNEASY lies the head that wears a crown," wrote Shakespeare. The same could be said today of the standard model of particle physics, our most successful description of the building blocks of matter and their interactions. The discovery of the Higgs boson in 2012 stands as the theory's crowning achievement, validating a prediction made nearly four decades ago and filling the model's last major gap. Yet we are as eager as ever to knock it from its throne, to discover the new physics that must surely supersede it. "The standard model is particle physics," says Nobel-prizewinning physicist Jack Steinberger. "But there are many unanswered questions that are extremely elusive at the moment."

Those questions include the nature of dark matter – the mysterious, invisible material thought to make up more than 80 per cent of the mass of the universe. Then there is dark energy, the stuff reckoned to be causing the universe's expansion to accelerate. In what must rank as our worst prediction, particle physics overestimates dark energy's magnitude by a factor of 10^{120} . The standard model also cannot explain how matter survived the big bang, or how gravity fits into the picture. It is riddled with so-called "free parameters", troublingly arbitrary numbers that have to be fed into the theory by hand, for example to set the strength of the interactions it describes.

Researchers had hoped that the Higgs would lead to the new physics that is needed

to explain away these difficulties. But with the Higgs behaving largely as expected so far, the real key to the kingdom beyond the standard model may lie with a different sort of particle: neutrinos.

Neutrinos hit the headlines in September 2011 when the OPERA experiment under the Gran Sasso mountain in Italy clocked them apparently travelling faster than the speed of light, an activity forbidden by Einstein's special theory of relativity. Six months later, the finding was traced to a glitch in the experiment. Even so, there is plenty more to say and learn about these beguiling particles.

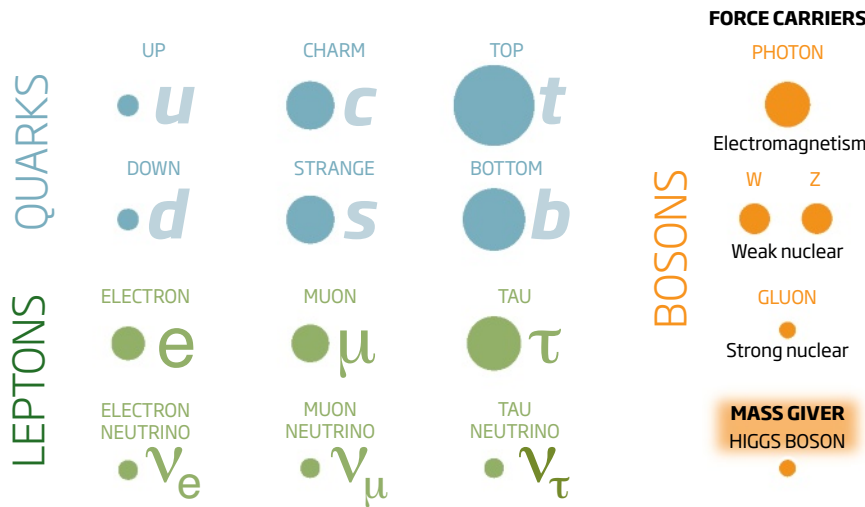
Ghostly, mysterious and antisocial, neutrinos rarely deign to interact with the world of common matter around them, and much of what we know about them lies outside the standard model. The three neutrinos we know about fit neatly enough. They pair with the electron and its two heavier cousins, the muon and the tau. A trio of antineutrinos also exists, which pair with the positively charged antiparticles of the electron, muon and tau to complete the extended lepton family (see chart, page 112). But at the outset, the standard model wrongly assumed neutrinos have no mass, and even now it cannot specify the masses they do have. It did not foresee their ability to shape-shift from one type into another, nor the fact that there might be more than three of them.

Many new theories hope to fill in those ➤



The standard model

Our best understanding of the building blocks of matter and the forces that glue them together



gaps, including grand unified theories, supersymmetry and string theory. One of them might gain traction by explaining why neutrinos are so very weird. Neutrinos themselves might in turn tell us which theory is on the right track.

Despite their aloof nature, neutrinos have a long history as problem-solving particles. Physicist Wolfgang Pauli conceived of them in 1930 in order to conserve energy and momentum in radioactive beta decays. More recently, neutrinos have moved to the forefront of our efforts to explain how matter came to dominate antimatter in our universe. "Neutrinos allow you to access another world for the simple reason that they are not so strongly interacting with us in the visible world," says theorist Patrick Huber at Virginia Tech in Blacksburg.

Flavour change

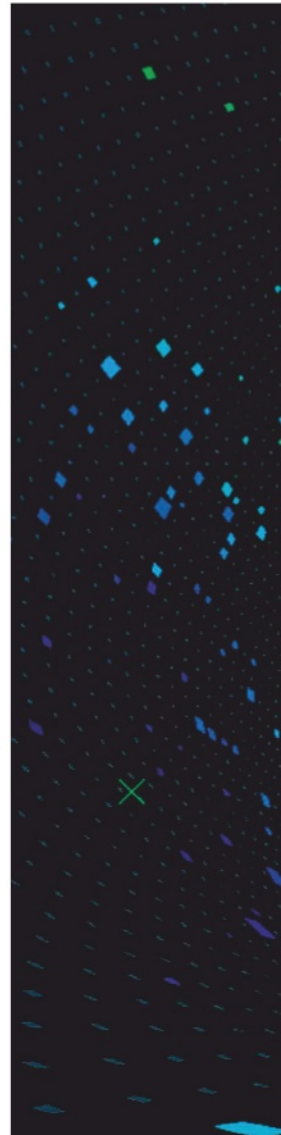
The first cracks in the standard model's description of neutrinos came over 15 years ago. Up until then, most physicists assumed – as did the theory – that neutrinos are massless. However in 1998, the Super-Kamiokande experiment in Japan proved that this wasn't the case. Neutrinos are emitted or absorbed with electron, muon or tau flavour, like single scoops of ice cream. Super-Kamiokande studied muon neutrinos from cosmic rays striking the atmosphere and found they could morph into electron neutrinos on their way through Earth. Other experiments investigating neutrinos created in nuclear reactors, particle accelerators and nuclear decay processes in the sun have confirmed that, however they start out, neutrinos shape-shift into a tutti-frutti mixture of flavours on their journey, with each scoop containing

a hint of all three. According to quantum mechanics, the only way such morphing can happen is if neutrinos have mass. Indeed, we now understand that each of the three neutrino flavours propagates through space as a different, constantly changing mixture.

That leaves us with a conundrum. "Neutrino mass tells us that the standard model needs to be extended, but it doesn't tell us how," says theoretical physicist Lawrence Krauss at Arizona State University in Tempe. In contrast, some grand unified theories – which go further by attempting to unite all the forces of nature except gravity – do predict neutrinos with mass, so pinning down the actual masses could tell theorists which theory to pursue. "There have been decades where people have speculated about grand theories which can explain the masses in various ways," says Joe Formaggio at the Massachusetts Institute of Technology, "but if you're going to come up with a theory that explains masses, you'd better have the masses."

Measuring the mass of an invisible particle that can sail unhindered through a slab of lead a light year thick is easier said than done. Catching neutrinos is a matter of patience, of watching long enough with big enough detectors until one eventually interacts. To do it, we have been stalking neutrinos at two radically different scales – the subatomic and the cosmic. Seventy years ago, Enrico Fermi envisioned measuring neutrino mass by measuring radioactive beta decays. In a typical beta decay, a neutron inside an atomic nucleus turns into a proton while spitting out an electron and an electron antineutrino. Although the antineutrino is undetectable directly, Fermi outlined how its mass could be inferred from the energy and momentum of the accompanying electron. Neutrinos,

A muon neutrino leaves its mark in the Super-Kamiokande detector



"With the Higgs boson behaving largely as expected, the key to new physics may instead lie with neutrinos"

however, are so light that it has been impossible to achieve the sensitivity needed. An exquisitely sensitive experiment being built at the Karlsruhe Institute of Technology in Germany, called KATRIN, may yet win the race to accomplish that in the next few years.

Meanwhile, the tightest limits on neutrino mass come from the cosmos: the particles' fingerprints can be found on the mix of elements created in the big bang and supernovae, on the expansion rate of the universe, on the cosmic microwave background (CMB), and on how matter coalesced into galaxies and galaxy clusters.

A combination of cosmological

measurements reveal that the sum of the three neutrino masses cannot exceed more than about 0.3 electronvolts (eV) – more than a million times smaller than the next lightest particle, the electron. "To me it's exhilarating that you can look at all the galaxies and clusters in the universe and detect the mass of this tiny particle," says Scott Dodelson, a cosmologist at Fermilab in Batavia, Illinois. Frank Close of the University of Oxford thinks that message should be taken to heart. "We don't appreciate the magic of what we're doing," he says. Analysis of observations of the CMB from the Planck space observatory should significantly hone our limits on

the sum of the neutrino masses.

Breaking that sum down into the masses of the individual mass states is made difficult by their constant shape-shifting. Measuring the shifts allows us to draw inferences, and an analysis of the best existing data puts the mass of the lightest state at about 0.05 eV.

That still leaves a puzzle. "Why it is that neutrinos are so anomalously light compared with everything else is bizarre," says Close. "It's as if they want to be nothing and yet weren't allowed to be."

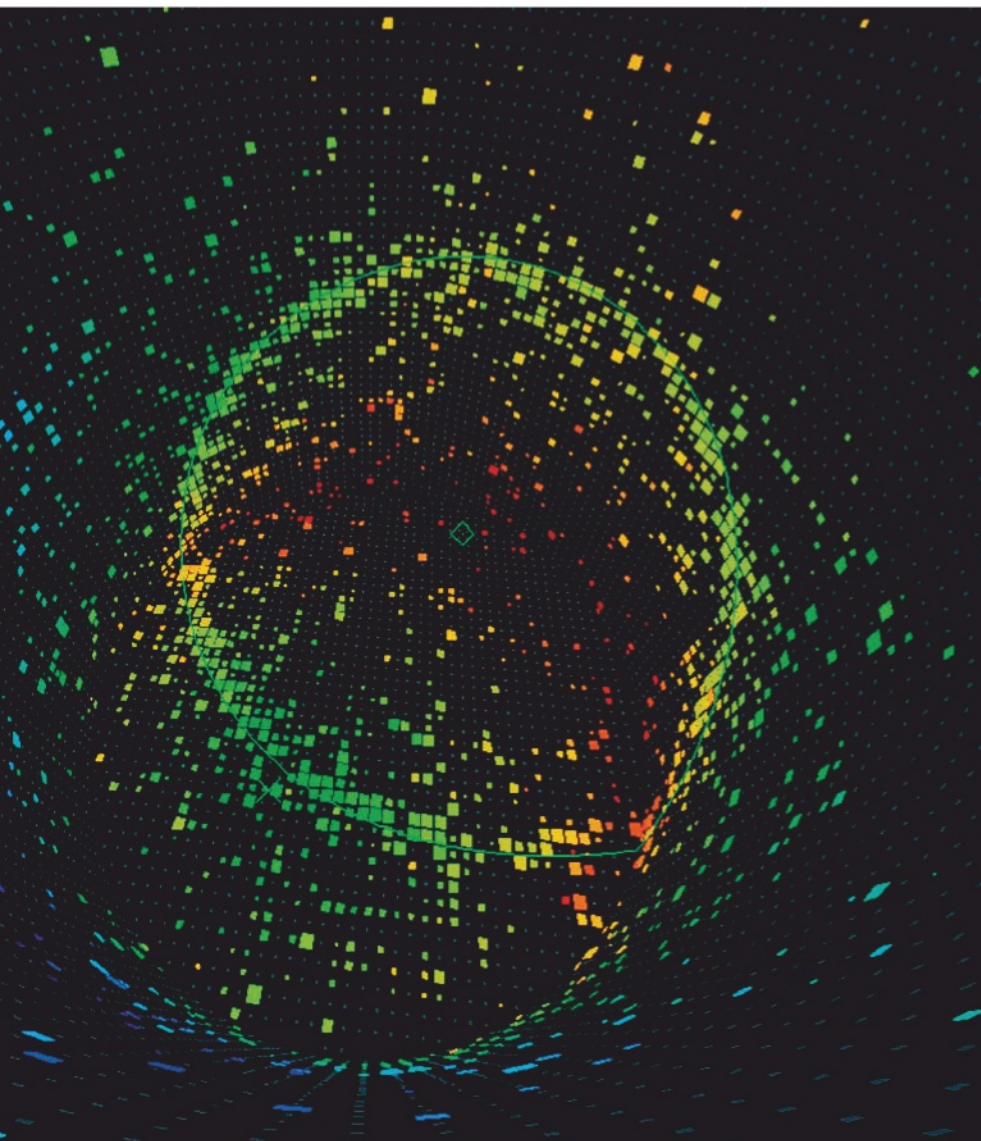
As if the three "normal" neutrinos were not antisocial enough, one theory suggests they may be shadowed by one or more "sterile" neutrinos. Unlike regular neutrinos, which feel the weak force inside nuclei and so occasionally interact with particles contained there, sterile neutrinos feel only gravity and so fail to interact with ordinary matter at all. Sterile neutrinos fascinate theorists since their discovery would break away from the standard model, and help explain not only dark matter but perhaps why there is matter at all. "They may well participate in forces beyond the standard model that we have not discovered yet," says theoretical physicist Boris Kayser of Fermilab.

Matter wins

Over the years, experiments have spun off a string of anomalies that point to one or more sterile neutrinos with a small mass of about 1 eV (see "Strange surplus", page 114). Predicted neither by the standard model nor by grand unified theories, their confirmation would hand researchers just the kind of new physics they are looking for.

The 2012 publication by an international group of almost 200 neutrino physicists of a "white paper" on sterile neutrinos reflects the interest they have stirred up. It describes some 21 experiments that are running, planned or proposed to try to track them down. "A large number of institutions are getting very excited about this," says Carlo Rubbia, a Nobel-prizewinning particle physicist at CERN. "We hope progress is coming fast."

Along with sterile neutrinos, researchers are stalking another prize – a difference between neutrinos and antineutrinos that could help explain why our universe is dominated by matter, and so why we are here to notice. According to our best understanding of cosmology and particle physics, matter and antimatter were created in equal amounts at the big bang. What followed was a



TOMASZ BARCZAK/SUPER-KAMIOKANDE COLLABORATION/SCIENCE PHOTO LIBRARY

"We have no good explanation for why the universe is made entirely of matter. It's a very embarrassing problem"

maelstrom of interactions, and in this melee matter and antimatter should have annihilated to leave nothing but a cosmos full of light. Clearly this hasn't happened. "We have no good explanation for why the universe is made entirely of matter," says Janet Conrad at MIT. "It's a very embarrassing problem."

"It's perhaps the most fundamental question we can ask about the universe, and neutrinos can provide a window into that question," says Alexandre Sousa at the University of Cincinnati.

That window is a theory called leptogenesis, and it relies on a phenomenon called CP violation. What this means is that if you look at a particle reaction, and then the same reaction viewed in a mirror and with particles swapped for their antiparticles, you will see the two reactions proceeding at slightly

different rates. It has been spotted in lab experiments with composite particles made up of quarks, but the imbalance seen there is not sufficient to explain why the antimatter created in the big bang vanished. The idea of leptogenesis is that in the first microseconds after the big bang, the young, hot universe contained extremely heavy, unstable sterile neutrinos that soon decayed, some into leptons and the remainder into their antimatter counterparts, but at unequal rates. This imbalance need only be tiny – one part in a billion. But it would mean that when the matter mopped up all the antimatter, enough leptons remained behind to eventually transform into the protons and neutrons that went on to form stars, galaxies and planets.

Heavy sterile neutrinos and their standard-model counterparts are thought to have been inextricably linked in the early universe:

according to a theoretical process known as the see-saw mechanism, neutrinos acquired their puzzlingly light masses by interacting with their heavyweight counterparts when the universe was extremely hot. If the picture of leptogenesis is true, we should see neutrinos and antineutrinos behaving in a slightly imbalanced way too.

So far, experimentalists have not uncovered any convincing neutrino CP anomalies. Fermilab's MINOS experiment created a buzz in 2010 when it found slight differences in the way that muon neutrinos and their antineutrino counterparts shape-shift as they travel over long distances, but by 2012, with more data, the difference disappeared.

Still, the prospects for glimpsing CP violation are good. In 2012, researchers at the Daya Bay Reactor Neutrino Experiment, based in southern China, measured a crucial parameter

STRANGE SURPLUS

It was a few flashes of light two decades ago that started the story of the biggest neutrino anomaly of them all. They occurred at the Liquid Scintillator Neutrino Detector (LSND) at Los Alamos National Laboratory in New Mexico, and each represented the passage of a neutrino through the detector's massive tank of mineral oil. Those flashes revealed that more muon antineutrinos than expected had changed into electron antineutrinos en route from a particle accelerator 30 metres away.

The leading explanation for the surplus is that on their way they briefly morph into undetectable "sterile" neutrinos, giving them another route to effect their transformation. By 1998, when LSND ended, the excess was still there and had reached a significance of 3.8 standard deviations – not enough to claim an outright discovery of sterile neutrinos, but sufficient to claim hints of them at work. "We were left with a very surprising result," says Bill Louis at Los Alamos, who worked on the experiment.

Still, the LSND anomaly would probably have faded into oblivion had it not been bolstered by a series of similar findings.

Researchers at Fermilab in Batavia, Illinois, built the MiniBooNE experiment to check LSND's results. It started by looking for muon neutrinos morphing into electron neutrinos, although at higher energies and

over a longer distance than LSND. Then it switched to antineutrinos like LSND. The details are complicated, but it too found hints that sterile neutrinos might exist.

A completely different experiment has also suggested the existence of sterile neutrinos. One of the early experiments to detect neutrinos streaming from the sun used tanks of gallium, which solar neutrinos could transmute into a detectable germanium isotope. Researchers calibrated their detectors using known radioactive sources. In two separate projects, based underground in Italy and Russia, detectors snared 15 per cent fewer neutrinos than expected from models of how many should have been produced – the so-called GALLEX and SAGE anomalies. Again, a likely explanation is that some neutrinos shape-shifted into an undetectable form.

BIG SPLASH

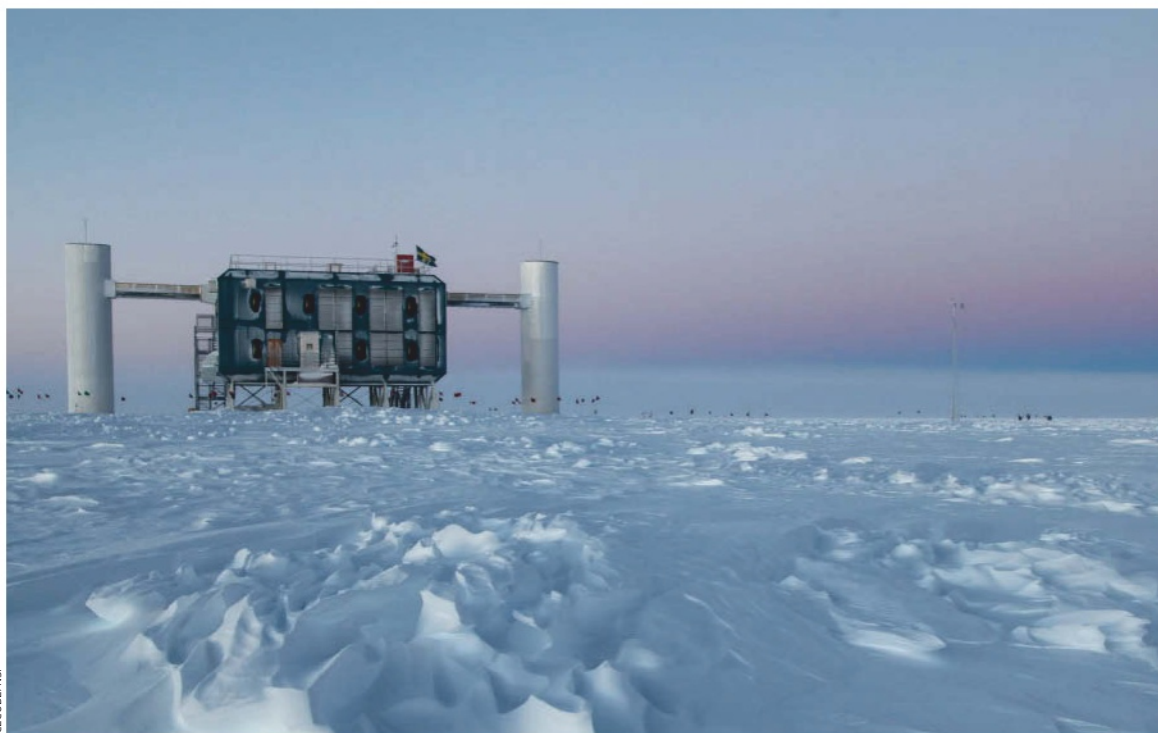
Then there are the anomalies at nuclear reactors. Improved calculations of the way nuclei capture neutrinos, and how many neutrinos nuclear reactors generate, indicate that several experiments over the past three decades should have found on average 7 per cent more neutrinos than they actually did. "When we discovered this anomaly, we were not looking for sterile neutrinos at all," says Thierry Lasserre, a

neutrino physicist at CEA, in Saclay, France. "It was a big surprise to us."

Louis checks off MiniBooNE, SAGE, GALLEX and the reactor anomalies. "All of those appear to be consistent with LSND," he says. "This has given additional incentive to look into sterile-neutrino models."

Janet Conrad at the Massachusetts Institute of Technology and her colleagues have published a very promising model which proposes three sterile neutrinos paralleling the three flavoured ones. The model explains most of the anomalies found close to neutrino sources. "You can't assume that there's just one sterile neutrino," Conrad says. "We put in three plus three and get a very good fit for both the disappearance and appearance data. We think that's going to be very big and splashy."

Lasserre proposes more experiments to settle the problem. He wants to insert an intense radioactive source into the heart of an existing detector. If light sterile neutrinos with a mass of about 1 electronvolt are produced by such a source, they should oscillate relatively fast into and out of detectable flavours. "You would see these beautiful oscillating patterns," says Lasserre. "If you manage to do this, either you find something or you are sure there is no sterile neutrino." He hopes to have a final answer within the next few years.



The South Pole is a fertile hunting ground for neutrinos

called θ_{13} , which describes how neutrinos change flavour. A low θ_{13} would have made CP violation hard to find, and zero would have ruled it out. To the researchers' delight, however, the value turned out to be surprisingly large, implying that future experiments have a good chance of finding CP violation.

"We now think we have the big picture," says André de Gouvêa, a theorist at Northwestern University in Evanston, Illinois. A first glimpse of the detail may come from Fermilab's Nova experiment, touted to have the best chance yet to detect neutrino CP violation. "It's the one experiment that can look at this over the next decade," says Sousa.

Even if neutrinos show CP violation, it is only part of the story. Leptogenesis only works if neutrinos, including the sterile variety, are so-called Majorana particles. This means that, unlike most other particles in the standard model, they are identical to their antiparticles and get their mass through the see-saw mechanism.

If this is indeed the case, we would expect to observe a process known as neutrinoless double-beta decay that the standard model frowns upon. In normal beta decay, a neutron changes into a proton and emits an electron and an electron antineutrino. Some nuclei can undergo two such decays at once, in which case we would expect two antineutrinos to be emitted. If these antineutrinos are identical to neutrinos, however, they will annihilate each other on emission, and the reaction will

produce just two protons and two electrons.

"Neutrinoless double-beta decay is the smoking gun that neutrinos are Majorana particles," says Alan Poon of Lawrence Berkeley National Laboratory in California. "It would give lots of tips to theorists on how to update the standard model, and it ties back to the very early universe – how we got more matter than antimatter."

Chasing the dream

Another allure of neutrinoless double-beta decay experiments is that the mass of the neutrino influences the reaction rate, allowing us to pin down this quantity too. "You get two very interesting pieces of physics – the mass of the lightest neutrino and the fact that neutrinos are Majorana particles," says Art McDonald, a particle astrophysicist at Queen's University in Kingston, Ontario, Canada.

So far, only one group claims to have seen neutrinoless double-beta decay, a Russian-German collaboration that first published their study of germanium decays in 2002. No other experiment has replicated their results. Findings from the Enriched Xenon Observatory, near Carlsbad, New Mexico, using a bath of liquid xenon, show that if neutrinoless double-beta decay exists at all, it is extremely rare – perhaps vanishingly so. Nevertheless, so great would be the prize of observing it that it remains the object of multiple research projects.

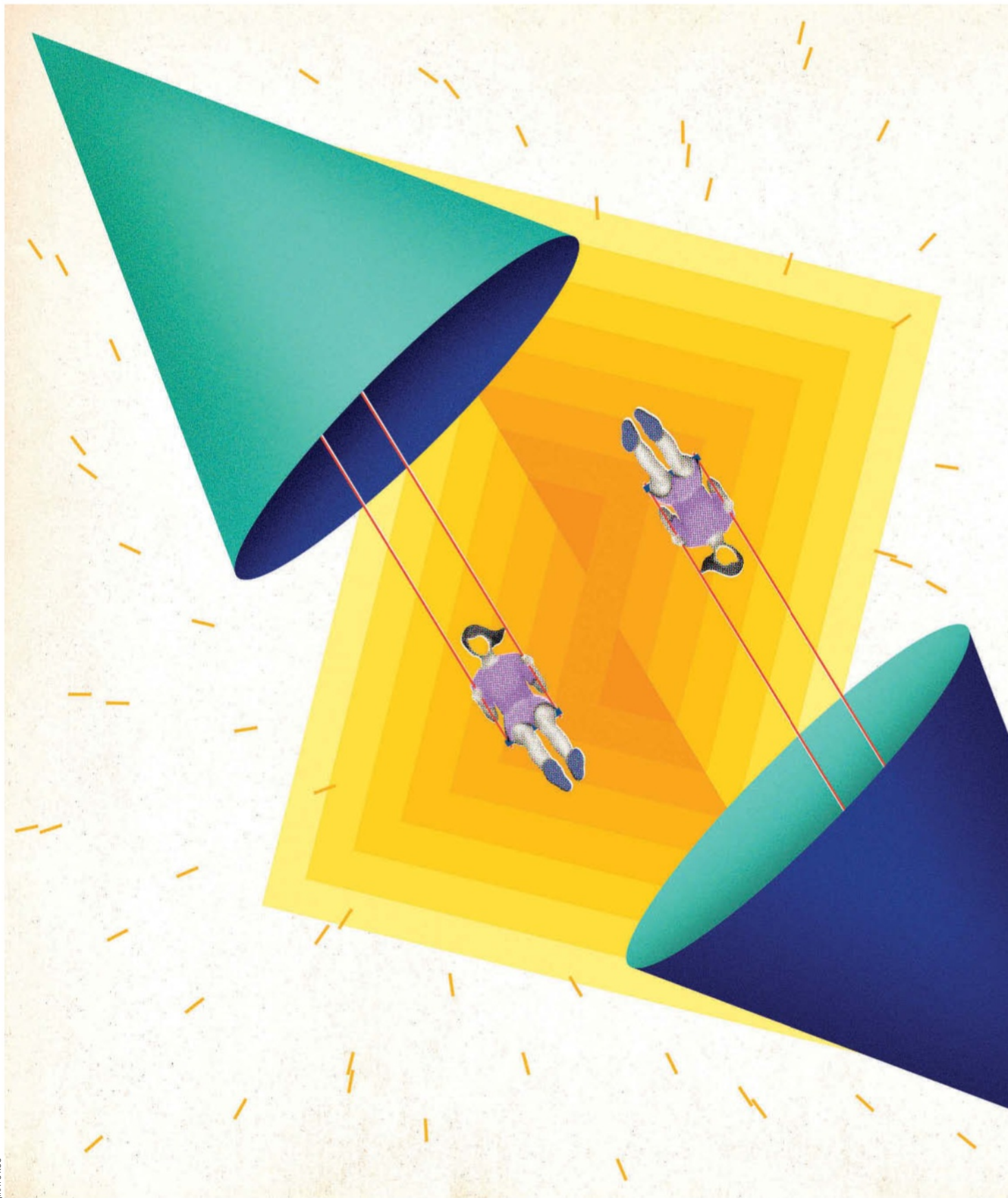
Many questions about neutrinos remain

open. Sheldon Glashow, a Nobel-prizewinning theorist at Harvard University, says what is needed are more and better experiments. "I don't think there's much to do until we have some experimental guidance," he says.

Francis Halzen, who heads the IceCube Neutrino Observatory, an experiment to measure cosmic neutrinos passing through Earth that is situated under the ice at the South Pole, agrees. "We chase new physics connected with neutrino oscillation. We may discover that neutrinos have non-standard-model interactions. We may discover there are sterile neutrinos mixing in with the three standard neutrinos," he says, "or something totally out of the blue."

The problem, they point out, is resources. Among the next experiments that have been proposed is the Long Baseline Neutrino Experiment, managed by Fermilab. This would be an intense neutrino beam fired hundreds of kilometres through Earth's mantle to a large detector weighing many thousands of tonnes. Another is the UK-to-Japan Neutrino Factory, which would create an intense beam of neutrinos and ping it to a detector on the other side of the world. Both would take decades to build and cost many billions of dollars.

It's worth the money and effort, says Rubbia. "This is one of the areas in which new discoveries are possible, but we don't know from which direction these discoveries will come. So we have to take a very courageous view to find out what's coming next." ■





Our hopes of finding an ultimate theory depend on upsetting a balance that Einstein cherished, says **Stuart Clark**

Differently equal

COINCIDENCE is not generally something scientists have much truck with. If two things are genuinely unrelated, there is little further of interest to be said. If the coincidence keeps turning up, however, there must be some deeper underlying link. Then it is the job of science to tease out what it is and so explain why there was no coincidence in the first place.

That makes it rather odd that a large chunk of modern physics is precariously balanced on a whopping coincidence.

This coincidence is essential to the way we view and define mass. It is so fundamental to the world's workings that most of us encounter its consequences every day without giving them another thought. Yet it has vexed some of the best minds in physics for centuries. Galileo and Newton grappled with it and ended up just accepting it, rather than understanding it. Einstein went one better: he declared it a principle of nature. He went on to use this "equivalence principle" as the fundament of his general theory of relativity, still our best stab at explaining the mysterious force of gravity.

But there is a problem. If we want to find some bigger, better theory that can unify gravity with the other forces that dictate the world's workings, the equivalence principle cannot stay. We must either unmask this coincidence – or radically rethink how physics can progress from here.

There are many versions of the equivalence principle, but all boil down to one idea: that the effects of gravitational fields are indistinguishable from the effects of accelerated motion. A thought experiment of Einstein's expresses it best. Imagine a person

standing inside an elevator on Earth. What keeps their feet firmly planted on the floor? The inexorable downward pull of gravity, of course. Now imagine the same person standing in the same elevator, but in empty space far from any gravitating object. In this case a rocket just so happens to be pushing the lift up in empty space with the same acceleration that Earth's gravity produces. The passenger will remain squarely on the lift floor in exactly the same way (see diagram, page 118).

How so, when there is no gravity involved? In this case, it is the person's inertia that is preventing them floating upwards. Inertia is the natural resistance of any body to acceleration – the same effect that pushes you back into your car seat when the driver puts their foot down.

The two elevator situations have a common property, mass. But the two masses come from very different places. One, gravitational mass, is something that responds to the pull of gravity, tending to accelerate a body in a gravitational field. The other, inertial mass, is the property of a body that opposes any acceleration.

Another way of stating the equivalence principle is to say that these two masses are always numerically exactly the same. The consequences of this coincidence are profound. If the two masses weren't the same, objects of different masses could fall to Earth at different rates, rather than all accelerating in the same way in a gravitational field. This "universality of freefall" was apocryphally first tested by Galileo dropping a bag of feathers and a bag of lead shot from the Leaning Tower of Pisa. In fact, the equality of ➤

An enigmatic equivalence

gravitational and inertial mass dictates all gravitational motion throughout the universe. If gravitational mass responded just a little bit more to gravity than inertial mass does to acceleration, for example, then planets would orbit their stars and stars orbit their galaxies just a little bit faster than they do.

Yet there is no obvious reason why this correspondence should be so. It was only by assuming it was that Einstein fully developed the strange contortions and contractions of time and space he had introduced in his special theory of relativity in 1905. What if a massive object such as a planet, Einstein wondered, squeezes the surrounding space into successively more compact volumes the closer you get to it? As something moved towards the planet's surface, it would then take less and less time to cross these compacted spaces: it would appear to accelerate.

The odd force

By 1916, this thought had guided Einstein to his general theory of relativity. What looks like gravity is just uniform motion through a progressively compacted space. And if there is no gravity, gravitational mass is fictitious too. The only mass at work in the universe is the one that gives a body its inertia. The coincidence behind equivalence disappears.

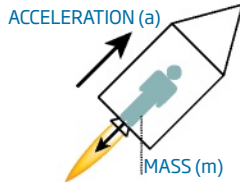
General relativity is, as far as we have tested it, peerlessly accurate, predicting the positions of celestial bodies and guiding our satellites with minute precision. Yet there is something odd about it that physicists don't like. All the other forces of nature are transmitted between bodies by physical, if ethereal, quantum particles. The electromagnetic force, for example, is transmitted between bodies with electrical charge by the exchange of the massless particles called photons. Outwardly, gravity works in exactly the same way. It looks like a duck, swims like a duck – but it can't quite be made to quack like a duck.

Attempts to make gravity quack with a quantum voice are the guiding thought behind string theory and other projects to construct an all-embracing "theory of everything". But if gravity is to be reborn as a real force, it needs something to latch on to, just as electromagnetism latches on to electric charge. It needs a gravitational mass that is separate and distinct from inertial mass.

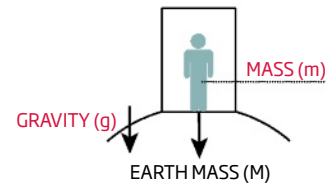
That means progress towards a theory of everything has an essential first step: slaying Einstein's holy cow. "Any theory of quantum gravity must violate the equivalence principle at some level," says Ben Gripaios, a theoretical

Einstein's equivalence principle states that the physics of acceleration and gravity work in exactly the same way. But there's no reason why that should be the case

Accelerate a rocket in gravity-free space and a body's **inertial mass** will resist the motion



The mutual attraction between **gravitational masses** is what keeps our feet on the ground



In all situations

$$\begin{aligned} \text{inertial mass} &\equiv \text{gravitational mass} \\ \text{acceleration} &\equiv \text{gravity} \end{aligned}$$

It's only by assuming the equivalence principle is true that we can explain that bodies at the same distance from Earth fall to the ground at the same rate

Newton's 2nd law of motion

Newton's gravitational law

$$\begin{aligned} \text{force} &= \text{mass} \times \text{acceleration} = \text{mass} \times \frac{\text{Gravitational constant} \times \text{Mass of Earth}}{r^2} \\ \text{If } m &= m \\ \text{then } a &= \frac{G M}{r^2} = g \end{aligned}$$

At a distance (r) from Earth's centre, the acceleration due to gravity (g) is ALWAYS THE SAME

physicist at the University of Cambridge.

How? One tried-and-tested method is to attempt to prove that the two masses aren't actually equivalent at all – just very, very close. Even the slightest sliver of a difference would mean that general relativity is built on an approximation and that a deeper, more precise theory must exist. "If someone finds a difference then we have made a major breakthrough," says Claus Lämmerzahl of the University of Bremen in Germany.

A way to do this is to continue on in the spirit of Galileo's Leaning Tower experiments,

"Gravity looks like a duck and swims like a duck – but it can't quite be made to quack like a duck"

testing the universality of free fall and other consequences of the equivalence principle in the hope of teasing out some tiny anomaly – so far with little success (see "Drop the subject", page 119).

Meanwhile, theorists are picking at a different thread. They point out that whether or not Einstein was right about there being no gravity, just inertia, no one has yet come up with a convincing explanation of inertia. "We do not yet know how to define it," says Gripaios. "We know it must be related closely to mass, but until we can define it precisely and know how to measure it, there can be no theory for it."

One thing's for sure: it doesn't all come from the Higgs field, feted as the giver of mass. Evidence for the existence of this field and its associated particle was presented by physicists sifting through the debris of particle collisions

at the Large Hadron Collider at CERN near Geneva, Switzerland, in 2012. But while the Higgs field is thought to give fundamental particles such as electrons and quarks their mass, when quarks combine into the heavier particles, protons and neutrons that make up the bulk of normal matter, the resulting mass is roughly a thousand times the summed mass of the constituent quarks. This extra mass comes not from the Higgs mechanism, but from the energy needed to keep the quarks together. Somehow, these two effects must combine and latch on to something else to create the property of a body's resistance to acceleration. "There is no way the Higgs alone can be some sort of mysterious ingredient that gives inertia," says Gripaos.

What then? One suggestion has its origins in work by Stephen Hawking in the 1970s. Ironically, it was motivated back then by a strict application of the equivalence principle. Hawking was investigating the properties of black holes, the unimaginably dense gravitating bodies whose existence is a central prediction of general relativity. He suggested that a black hole should be an apparent source of radiation, because pairs of quantum particles that constantly pop up in space would become separated close to a black hole, with one being sucked in and the other spat out. That led the Canadian physicist William Unruh and others to suggest that, if gravitation and acceleration really are one and the same thing, similar emissions should be a feature of any body accelerating in a vacuum.

Nothing doing

Like Hawking's radiation, Unruh's has never been unambiguously detected. The accelerations necessary to achieve a measurable effect in a lab are generally too high, although some argue the effect has been seen with electrons accelerated in the high magnetic fields of particle accelerators.

A decade or so on from Unruh's original work, astrophysicist Bernard Haisch, then at the Max Planck Institute for Extraterrestrial Physics in Garching, Germany, and electrical engineer Alfonso Rueda of California State University in Long Beach were playing with a similar idea when they realised the vacuum's interaction with an accelerating body would not just occur on its surface, but permeate its entire volume. That could produce a force that acts in the opposite direction to the body's movement. They originally likened it to the way in which charged particles moving through a magnetic field experience a force –

DROP THE SUBJECT

Looming 146 metres over the north German plain like a great white rocket, it's hard to ignore the University of Bremen's "drop tower" (picture, right). Inaugurated in 1990 as part of the Center of Applied Space Technology and Microgravity (ZARM), it provides up to 9.3 seconds of free fall in which to conduct experiments. So far tests of rubidium and potassium atoms have provided no deviation from the behaviour predicted by the equivalence principle (see main story). The atoms have been found to fall at the same rate to accuracies of 11 decimal places.

At the University of Washington in Seattle, meanwhile, Eric Adelberger and his "Eöt-Wash" team use a high-tech set of scales known as a torsion balance to compare the motions of standard masses made of different elements, including copper, beryllium, aluminium and silicon. They hold the record for test accuracy, with no violations of the equivalence principle to 13 decimal places.

At some point, however, these earthbound experiments are going to hit a brick wall. "It is getting harder to make the instruments better," says Adelberger. Working somewhere where gravity is a lot smaller would make any deviations from equivalence a lot easier to spot.



The French-led MICROSCOPE mission due to launch in 2016 will do just that, testing the motions of masses of platinum and iridium in the microgravity conditions of space. "MICROSCOPE will achieve an accuracy 100 times better than a laboratory on Earth," says Claus Lämmerzahl of ZARM.

His organisation is testing the satellite's accelerometers in their drop tower, and also developing the software needed to analyse the satellite's final results. An even more sensitive mission, the Space-Time Explorer and Quantum Equivalence Principle Space Test, is currently being evaluated by the European Space Agency.



If objects fall at different rates under gravity, the equivalence principle is broken



JIRAYU KOO

the Lorentz force – that affects their motion. In this case there were electromagnetic interactions with the quantum vacuum. “It appears to be exactly what you need for inertia,” says Haisch.

Anomalous accelerations

Mike McCulloch of the University of Plymouth, UK, thinks such interactions are also just what you need to break the equivalence principle. One prediction made of Unruh radiation is that, like the rays emitted from a hot body, it comes in a spectrum of many different wavelengths. For very small accelerations, the temperature of the radiation that a body “sees” from the vacuum is low, and dominated by very long wavelengths. Make the acceleration very small indeed, and some of these wavelengths become longer than the size of the observable universe, effectively cutting them off.

In this case, according to calculations McCulloch did in 2007, originally to explain the seemingly anomalous accelerations of the Pioneer spacecraft as they crossed the solar system, the total amount of Unruh radiation experienced by a body would drop, and it would feel less of an opposing force. Its inertia would thus fall, making it easier to move than

Newton’s standard laws of motion dictate – and cutting the connection with gravitational mass.

The problem with this idea is testing it. In the high-gravity environment of Earth, accelerations small enough for the effect to be observed would not be easy to manufacture. But its effects might well be seen in a low-gravity environment such as that found at the edge of a galaxy. Indeed, looking at the anomalous motions of most spiral galaxies, McCulloch suggests this mechanism could also explain another enduring cosmic mystery – that of dark matter (see “Dark inertia”, above).

It’s fair to say such ideas have not set the world alight. When Haisch and Rueda came up with their mechanism, NASA was sufficiently impressed to fund further study and the duo also attracted some \$2 million in private investment. But the lack of testable predictions of how the effect might manifest itself led the money and interest to dry up.

Nevertheless, a traditionalist such as Lämmerzahl thinks we should not dismiss the idea out of hand. “Even though I follow more the ideas of string theory, these ideas of vacuum interactions are not nonsense,” he says. “We need to look at them seriously and decide whether they give us new ways to test the equivalence principle.”

One proposal to do that was made in

DARK INERTIA

In the 1930s, we noticed that galaxies spinning around other galaxies were not moving as Newton’s and Einstein’s laws of gravity dictated. A few decades later, something similar was observed of the rotation of individual spiral galaxies. It was almost as if some invisible matter was whirling the matter we could see around faster.

That idea has now become mainstream: standard cosmology textbooks will tell you that “dark matter” outweighs normal matter by a factor of 5 to 1. Yet despite particle physicists supplying an almost endless list of hypothetical particles that might fit the bill, to date none has been definitively detected.

An alternative first championed in the 1980s by Mordehai Milgrom, then at Princeton University, is that gravity must somehow be modified at a galaxy’s edges. This could be explained if there was a drop in inertial mass without a drop in gravitational mass for stars experiencing the ultra-low accelerations found at the outskirts of galaxies. This would naturally make them move faster. If vacuum interactions can really bring this about (see main story), they could be just the ticket to mimic dark matter.

2010 by a trio of Brazilian astronomers led by Vitorio De Lorenci of the Federal University of Itajubá. They suggested using a spinning disc to cancel out the accelerations produced by Earth’s rotation and its movement through space. At minuscule accelerations, the disc’s inertia would drop, meaning it would spin faster than expected from Newton’s laws. Despite a relatively modest cost, however, no money has yet been forthcoming to fund the experiment.

And so the deadlock remains until someone delivers either an experiment that exposes the equivalence principle as a sham, or a theoretical idea that shows why it must be just so.

But if in the end gravitational mass is indeed just inertial mass in another guise – whatever inertial mass is – then it will be the quantum theories of gravity, including string theories, that will find themselves laid upon the sacrificial altar. Paths to a theory of everything will become even more winding. If gravity is not a force, but truly an illusion that springs from the warping of space, as described by general relativity, we will have to look more closely to understand at a basic level what makes that warping come about.

Just a coincidence? This is one that science is not finding so easy to dismiss. ■

Develop and share your passion for biology



Join

The Society of Biology represents, supports and engages with anyone who has an interest in the life sciences, and offers membership grades to suit all levels of expertise and interest.

Email us for more information, quoting ref: NS314, and you'll **save 25% on your first year of membership.**



membership@societyofbiology.org

or visit www.societyofbiology.org/join

Could a monstrous, unseen mathematical entity be pulling the strings of the cosmos, asks Matthew Chalmers

IF YOU'VE ever tried counting yourself to sleep, it's unlikely you did it using the square roots of sheep. The square root of a sheep is not something that seems to make much sense. You could, in theory, perform all sorts of arithmetical operations with them: add them, subtract them, multiply them. But it is hard to see why you would want to.

All the odder, then, that this is exactly what physicists do to make sense of reality. Except not with sheep. Their basic numerical building block is a similarly nonsensical concept: the square root of minus 1.

This is not a "real" number you can count and measure stuff with. You can't work out whether it's divisible by 2, or less than 10. Yet it is there, everywhere, in the mathematics of our most successful – and supremely bamboozling – theory of the world: quantum theory.

This is a problem, says respected theoretical physicist Bill Wootters of Williams College in Williamstown, Massachusetts – a problem that might be preventing us getting to grips with quantum theory's mysteries. And he has a solution, albeit one with a price. We can make quantum mechanics work with real numbers, but only if we propose the existence of an entity that makes even Wootters blanch: a universal "bit" of information that interacts with everything else in reality, dictating its quantum behaviour.

What form this "u-bit" might take physically, or where it resides, no one can yet tell. But if it exists, its meddling could not only bring a new understanding of quantum theory, but also explain why super-powerful quantum computers can never be made to work. It would be a truly revolutionary insight. Is it for real?

The square root of minus 1, also known as the imaginary unit, i , has been lurking in

mathematics since the 16th century at least, when it popped up as geometers were solving equations such as those with an x^2 or x^3 term in them. Since then, the imaginary unit and its offspring, two-dimensional "complex" numbers incorporating both real and imaginary elements, have wormed their way into many parts of mathematics, despite their lack of an obvious connection to the numbers we conventionally use to describe things around us (see "Complex stuff", page 125). In geometry they appear in trigonometric equations, and in physics they provide a neat way to describe rotations and oscillations. Electrical engineers use them routinely in designing alternating-current circuits, and they are handy for describing light and sound waves, too.

But things suddenly got a lot more convoluted with the advent of quantum theory. "Complex numbers had been used in physics before quantum mechanics, but always as a kind of algebraic trick to make the math easier," says Benjamin Schumacher of Kenyon College in Gambier, Ohio.

Quantum complications

Not so in quantum mechanics. This theory evolved a century ago from a hotchpotch of ideas about the subatomic world. Central to it is the idea that microscopic matter has characteristics of both a particle and a wave at the same time. This is the root of the theory's infamous assaults on our intuition. It's what allows, for example, a seemingly localised particle to be in two places at once.

And it turns out that two-dimensional complex numbers are exactly what you need to describe this fuzzy, smeared world. Within quantum theory, things like electrons and ➤

Reality bits



photons are represented by “wave functions” that completely describe all the many possible states of a single particle. These multiple personalities are depicted by a series of complex numbers within the wave function that describe the probability that a particle has a particular property such as a certain location or momentum. Whereas alternative real-number descriptions for something like a light wave in the classical world are readily available, purely real mathematics simply does not supply the tools required to paint the dual wave-particle picture.

Hidden complexity

The odd thing is, though, we never see all that quantum complexity directly. The quantum weirdness locked up in the wave function “collapses” into a single real number when you attempt to measure something: a particle is always found at a single location, for example, or moving with a certain speed. Mathematically, the first thing you do when comparing a quantum prediction with reality is an operation akin to squaring the wave function, allowing you to get rid of all the i 's and arrive at a real-number probability. If there is more than one way for a thing to end up with, say, a particular location, you add up all the complex number representations for each different way, and then square the sum.

The fact that this was a rather odd way to go about things first struck Wootters over 30 years ago, when he was writing his PhD thesis. In the “real” world, the probability of rolling a 10 from a pair of dice is $3/36$, because there are three ways to produce a 10 from the total of 36 possible outcomes. We add probabilities, we don't add them and square them. “This procedure would be strange even if the square roots were real,” says Wootters.

The involvement of complex numbers only makes things worse. Complex numbers are two-dimensional with real and imaginary parts, whereas the probabilities of observable reality are only one-dimensional – purely real. That implies some of the information stored in the complex numbers disappears every time we make a quantum measurement. This is very unlike the macroscopic, classical world, where there is at least in theory a perfect flow of information from the past to the future: given complete information about the properties and position of a pair of dice, for example, we can predict how they will fall.

Replace those complex square roots with real square roots, Wootters realised, and you can at least stem this troubling information



NATALIE NICKLIN

“Purely real mathematics simply does not supply the tools to paint the complete quantum picture”

haemorrhage. “Real square roots would be comprehensible – not a strange thing – if nature is interested in a strong link between past and future,” he says.

As it turned out, he wasn't the first to try to make quantum theory real. In the 1960s, Swiss physicist Ernst Stueckelberg had reformulated quantum mechanics using only real numbers. But when he tried to impose the fundamental quantum principle of uncertainty – the idea that we cannot determine quantities such as the position and momentum of a particle simultaneously with full accuracy – he found that he still needed something to play the role of the imaginary unit.

Back in 1980, this was all a shrug-shoulders affair for Wootters. He had other fish to fry, namely, helping to found a new branch of physics, quantum information theory. This

portrays quantum mechanics in terms of the information it can encode, and has allowed physicists to take advantage of fundamentally quantum processes such as teleportation and entanglement between particles to transport information more efficiently and securely than is possible by conventional, classical means. It also lies behind current attempts to build super-powerful quantum computers. Wootters's role as one of the founding fathers of the field has made him one of the most-cited physicists in the world.

It was an invitation in 2009 to give a talk in Vienna, Austria, that prompted him to return to the problematic role of i in quantum theory. Could all the quantum information theory developed in the meantime provide any new resolution?

By October 2012, working with his students Antoniya Aleksandrova and Victoria Borish, Wootters had progressed far enough to think it could. Conventional quantum theory talks of information in terms of qubits, probabilistic versions of the traditional binary bits that classical computers crunch. Wootters and his colleagues were able to replace these qubits with real-number equivalents, and so capture all the weird correlations and uncertainties of conventional quantum theory without an

i in sight. While he was about it, in 2013 Wootters also published his 30-year-old insight that a real-number quantum theory could solve the problem of imperfect information flow during quantum measurements.

All this came with a huge sting in the tail, however. Like Stueckelberg's earlier attempt, this theory also needs an extra element to play the part of i . It turns out to be a monster: a physical entity Wootters dubs the u-bit.

A u-bit is a master bit: an entity that interacts with all the other bits describing stuff in the universe. Mathematically, this omnipresent conduit of information is represented by a vector on an ordinary, real two-dimensional plane. What it represents physically, no one, least of all Wootters, can tell, but by entangling itself with everything else in the universe, it is sufficient to replace every single complex number in quantum theory. The mathematical description of the u-bit supplies one further, slightly whimsical clue to its physical identity: whatever and wherever it is, it must be rotating quite fast.

It all sounds rather like a mistimed April fool, but that might just be the idea's strength, says Schumacher. "I think a good paper in fundamental physics shares some characteristics with a good joke: it has an unexpected take on a familiar idea, and yet in retrospect it has a certain screwball inevitability. By that standard, Bill's u-bit theory is a very good joke."

Matt Leifer of the Perimeter Institute in Waterloo, Ontario, Canada, is more sceptical, pointing out that the nature of the u-bit makes it very difficult to test the theory. "It's not

"I think a good paper in fundamental physics shares some characteristics with a good joke"

yet clear whether it is something that can be manifested at a particular location such that you could find it in a detector like we did the Higgs boson," he says.

But the u-bit's influence might be felt indirectly. Since it interacts with everything in the universe, it can rough up even a supposedly isolated quantum system, making it "decohere" and lose its vital quantum properties in a way not predicted by standard quantum theory. That could be bad news for the quantum computers built on the back of quantum information theory, which are very sensitive to environmental disturbance. With a u-bit lurking off-stage, physicists could try all they like to seal off such a device from the outside world, for example by putting it inside a huge fridge to cool down its interactions, but there would be nothing that could shield it from decoherence induced by the u-bit.

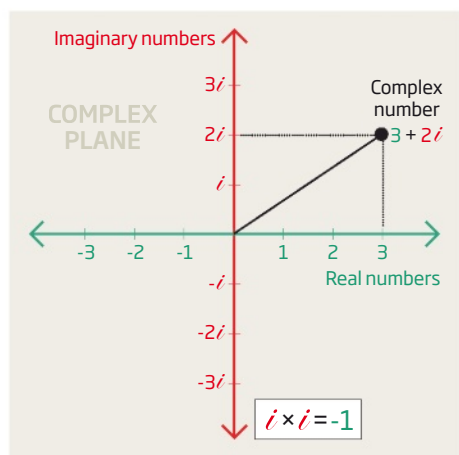
The length of time a conventional qubit could stave off decoherence is governed by the strength of the u-bit's interaction and its rate of rotation. An experiment carried out in 2011 by researchers at the National Institute of Standards and Technology in Boulder, Colorado, showed that qubits made out of trapped ions can remain quantum-coherent for a period of several seconds, a result that would suggest

the u-bit interaction is weak, if it exists at all. "The theory is definitely sensible, but it is currently not clear whether the u-bit is really realised in nature," says Markus Müller of Heidelberg University in Germany.

For an information theorist like Wootters, it is easy to see the attraction of an entity that does away with the profligate waste of information seemingly hard-wired into the mathematics of quantum theory. But he is the first to admit that the u-bit is no panacea for quantum ills. It doesn't itself explain what wave-particle duality is, or why a particle can be in two places at once. Something like entanglement is actually more pervasive in the new formulation, because the u-bit can be fully entangled with any number of objects. In conventional, complex-number quantum mechanics, full-on entanglement is limited to two objects.

But even if the real-number reformulation just turns out to be a new angle from which to view quantum theory's mysteries, that could be valuable. Despite quantum theory's enormous success in agreeing with experiment, its conflict with our intuitions has left us casting round for a "narrative" – a compelling explanation of why things should be the way quantum theory suggests they are. Often, those interpretations have been impeded by dwelling too much on one aspect of the theory, says Schumacher. "There has been an amazing amount of nonsense written about quantum theory because people used to understand it only as a theory about waves, for instance." That makes one insight from Wootters's work – that we should look at the information-carrying capacity of a quantum state to understand the physics behind it – an important one, says Müller.

As for the question of quantum theory's irreality, perhaps we have just to learn to love i . After all, it is not just in quantum mechanics where its influence is felt. Complex numbers are also increasingly vital in describing optical waveguides, transitions between different states of matter and many other aspects of classical physics. "People always thought of complex numbers just as a tool, but increasingly we are seeing that there is something more to them," says mathematician Dorje Brody of Brunel University in London. History tells us we have come to accept the reality of other mathematical concepts such as zero and the negative numbers only after long tussles. Perhaps "imaginary" was an unfortunate choice of words when we came to name i . Unless, of course, instead of an i , a u-bit lurks, ubiquitous – and unseen. ■



Complex numbers, containing both **real** and **imaginary** elements, lie somewhere on a 2D "complex plane"

COMPLEX STUFF

The square root of minus 1 cannot lie anywhere on the standard number line stretching from negative to positive infinity. For minus 1 to have a square root implies the existence of a number that, multiplied by itself, makes a negative number. But mathematical logic dictates that any two real numbers, if both negative or positive, produce a positive number when multiplied together.

So i is given a second line all of its own at right angles to the standard number line. As the "imaginary unit", it forms the basis of a second numerical dimension, and with it a plane of numbers that have both "real" and "imaginary" parts – the space of complex numbers (see left).

THE GRAND FINALE

How will the universe end? Will it end at all?
Stephen Battersby explores our ultimate fate

THE future ain't what it used to be. Cosmologists were once confident they knew how the universe would end: it would just fade away. An ever colder, ever dimmer cosmos would slowly wind down until there were only cinders where the stars once shone. But that's history.

Today's science suggests many different possible futures, from violent crunches, rips or snaps to bizarre trips. The slow drift into darkness is still a contender, but fear not: it could be a lot more interesting than you might think – imagine the cosmos filled with giant diamonds.

Why this wealth of possibilities? Until relatively recently, the dominant force in the universe was thought to be gravity, and that meant there were only two options. Either the universe was dense enough for gravity to halt the expansion from the big bang and pull everything back together in a big crunch, or else it wasn't, in which case the expansion would carry on forever.

Most cosmologists thought endless expansion the more likely. Then, in 1998, they got a shock: the expansion isn't slowing down at all, but accelerating. That seemed to seal our fate, condemning the universe to a headlong rush to oblivion.

But researchers have now realised this gloomy outlook was premature, because no one knows what is causing the acceleration. We call it dark energy, but its origin and nature are a mystery. So how can anyone say what it is going to do in the future?

Astronomers do at least agree on what will happen to our neighbourhood. In about 6 billion years the sun will swell to become a red giant, boiling the oceans away and possibly swallowing the Earth. Then it will exhaust its

nuclear fuel and shrink into a white dwarf.

Our galaxy is in for a rough time too. We are heading towards another, larger spiral called Andromeda, and the Milky Way could collide with it in as little as 3 billion years, ultimately merging into one great elliptical galaxy.

If we could look out from this merged galaxy we would see other galaxies moving away, dragged ever faster by dark energy. But for how long? That depends on the nature of dark energy. For instance, it could decrease with time, or even become negative.

If that happens, the consequences will be extreme. Expansion will eventually halt and then reverse, so that galaxies rush back towards each other and start colliding at ferocious speeds. Eventually, everything will be crushed together in a big crunch, like the big bang in reverse.

An even more grisly end could be in store. It is possible that dark energy could become a stronger "phantom energy" causing space to expand ever faster and setting up a devastating positive-feedback loop called the big rip.

If this is right, then the end could come in as little as 40 billion years. It would be perhaps the most watchable doomsday that cosmologists have imagined, not entirely unlike the spectacle laid on in Douglas Adams's *The Restaurant at the End of the Universe*. Stars and planets will explode and eventually even molecules, atoms and their constituent subatomic particles will be torn asunder.

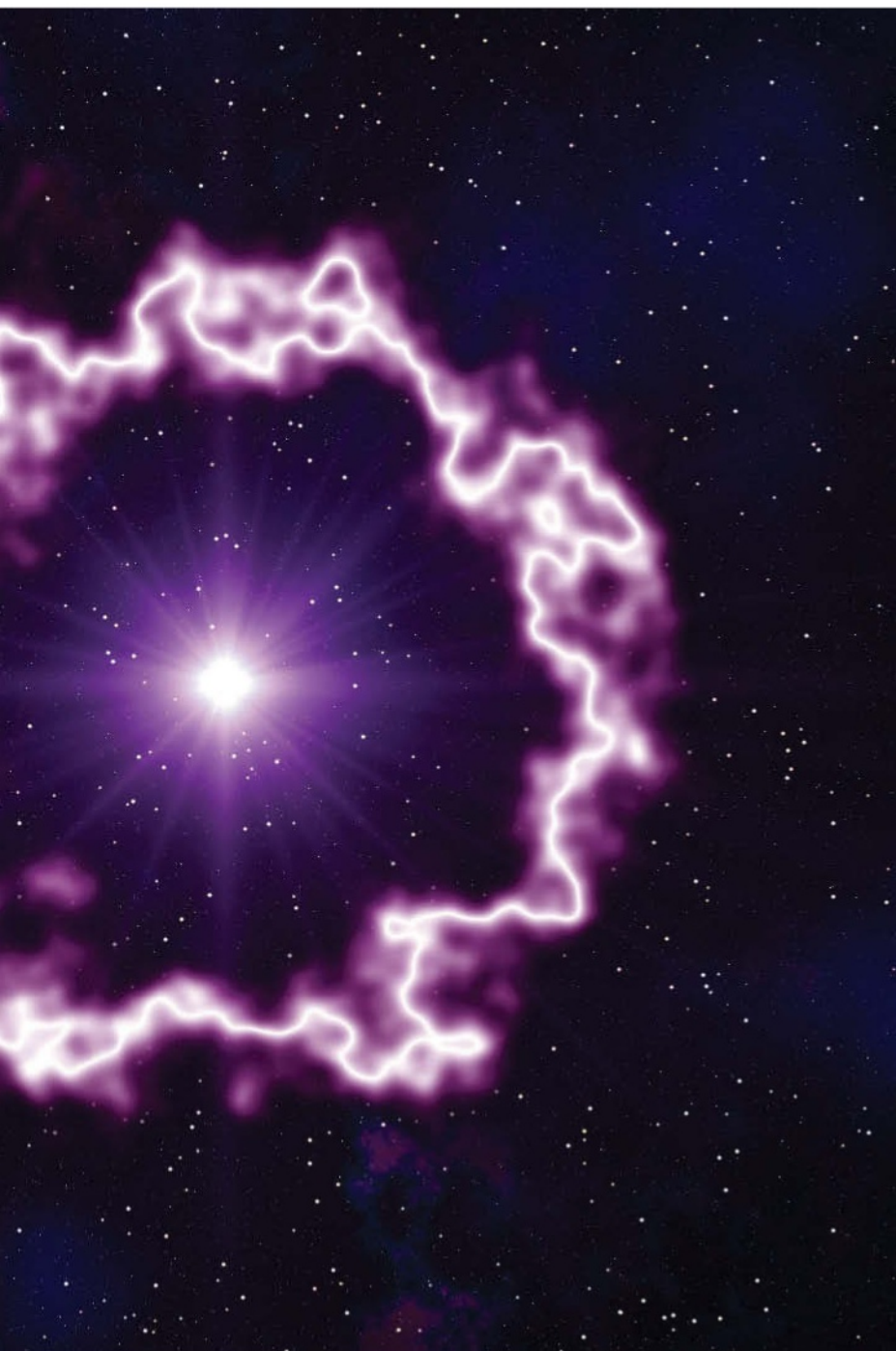
Phantom energy could have a different outcome, however, if our universe contains even a single wormhole. These tunnels in space-time would feed on phantom energy and grow. A phantom-fed wormhole could

grow large enough to swallow the whole universe – a scenario called the big trip. It is not clear where it would take us.

But there doesn't have to be anything changeable about dark energy. Many cosmologists think that it will simply stay constant. If so, our universe will steer carefully between crunch and rip.

Such a middle-of-the-road future may yet have a radical finale. According to quantum mechanics, the total amount of information in the universe should be constant. If space





keeps expanding, that might make things uncomfortable. The universe might eventually become pixelated, with information spread too thinly to support familiar physics. Everything would disintegrate in an event called the big snap.

If we avoid the big snap, then ordinary dark energy could lead to a long and lonely future. Acceleration will soon steal most of the universe away as the increasing expansion of space carries other galaxies beyond our cosmic horizon.

ROGER HARRIS/CFR

“Today’s science suggests many different possible futures, from violent crunches, rips or snaps to bizarre trips”

Then we will be all alone. After about a hundred trillion years, all nuclear-powered stars will have gone out. A little faint infrared radiation will come from stars called brown dwarfs. Other stars will be reduced to dense, dead remnants – black holes, neutron stars and ageing white dwarfs, slowly dimming to black. Our sun will become one of these black dwarfs: a single crystal of carbon, like an ultra-dense diamond. If there is a remnant of the Earth, then for a while it might trail after the black dwarf that was once our sun. But the system will slowly lose angular momentum, and Earth’s cinder will eventually spiral in to hit the sun’s.

Occasional flares will lift the gloom when brown dwarfs collide to form a new star or a black hole shreds a stellar carcass. Once in a trillion years, black dwarfs will collide and explode.

In about a hundred quintillion (10^{20}) years, our observable universe will be reduced to a diaspora of dead stars, loosely centred on a massive black hole.

Meanwhile, dark energy will still be at work. Each star will see all its old companions disappear over the horizon one by one. Our black dwarf will be in a universe of its own.

After that, it gets a lot more speculative, but here’s the best guess. Particle physicists suspect that protons are unstable and will probably only last a maximum of 10^{45} years. As protons decay into quarks, all the black dwarfs, neutron stars and planets will crumble away, leaving behind nothing but loose photons, neutrinos, electrons and positrons. Even black holes will eventually evaporate.

And then? Dark energy continues working. One day, every single particle will find itself alone inside its own horizon.

There is another possibility. In the cosmological models, some kinds of dark energy gradually fade, but never become negative. These scenarios would give us a chilly future where no stars shine and all solid bodies dissipate into a cloud of fundamental particles. However, the acceleration will eventually tail off, so it won’t isolate every particle. Particles could still interact, albeit at a glacial rate, and some kind of chilly life might just be able to cling to existence.

It may seem like a poor set of options – to be crushed, ripped apart or evaporated away. But in fact none of these scenarios need be terminal. All have a get-out clause that allows a new universe to be reborn from the ashes of the old one. So don’t think of the end of the universe as *the* end, but as the potential start of a new beginning. ■

COMING SOON

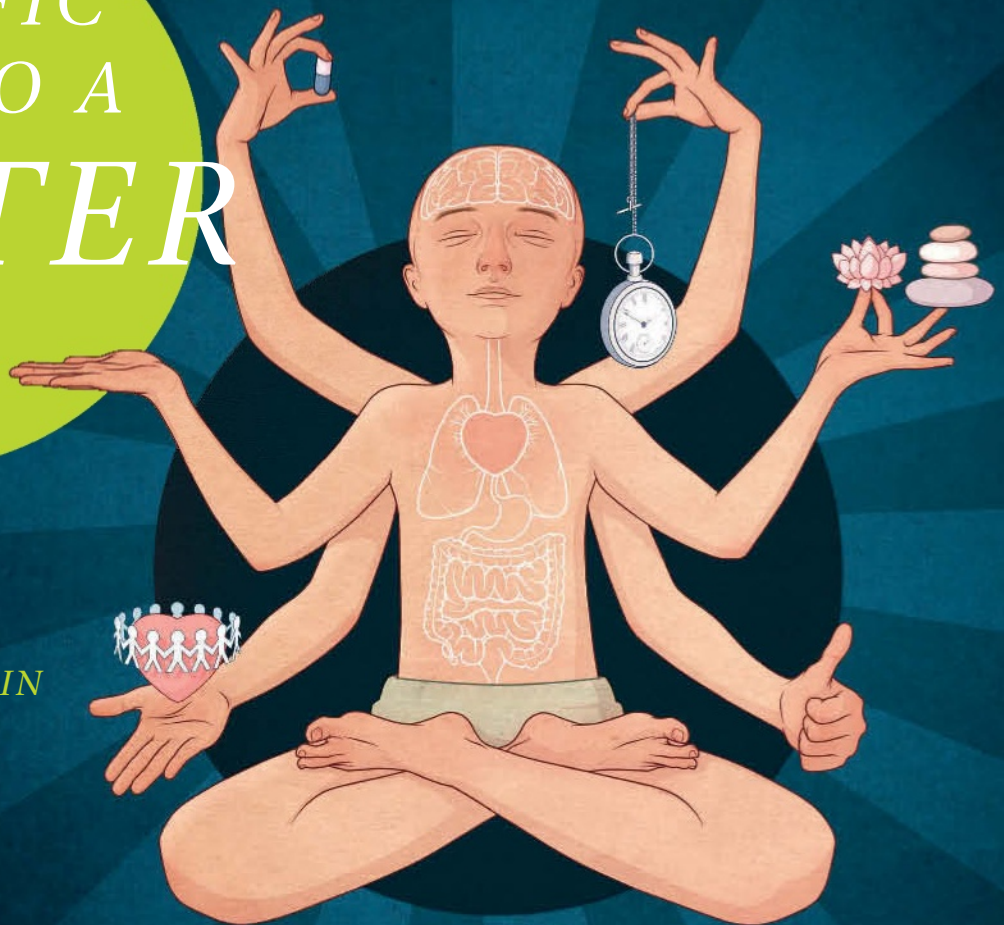
NewScientist

THE COLLECTION

ISSUE THREE

THE SCIENTIFIC GUIDE TO A BETTER YOU

HOW TO
STAY HEALTHY
SHARPEN YOUR BRAIN
EXERCISE WELL
EAT RIGHT
LIVE LONGER



ON SALE 13 AUGUST

If you missed the first issue of *New Scientist: The Collection*, visit newscientist.com/TheCollection to obtain your copy

It doesn't end here

Awe-inspiring. Enigmatic.
And unfathomably huge.

For thousands of years we have tried to unravel the mysteries of the universe - but despite uncovering many secrets, a tantalising amount remains unknown.

So what *do* we know?

Since 1956, one year before the first artificial satellite, *New Scientist* has been taking people to the limits of our understanding of the cosmos.

Subscribe today to hear every exciting development as they happen, as we close in on understanding the universe.

SPECIAL SUBSCRIPTION OFFER

Subscribe and save
Visit tinyurl.com/NSTheUnknownUniverse
or call +44 (0) 844 543 80 70 and quote 6714



NewScientist

ultra

Vorsprung durch Technik

Audi



Who says an efficient car can't be a desirable car? Our new A6 ultra models are available with all the features you really want like sports suspension, sports seats, S line styling and 20" alloy wheels, yet still return an ultra low emission rating.

The new Audi A6 ultra, with CO₂ from 114g/km.



The new Audi A6 ultra.



And its CO₂ emission rating.

Official fuel consumption figures for the Audi A6 2.0 TDI ultra Black Edition S Tronic in mpg (l/100km): Urban 55.4 (5.1), Extra Urban 72.4 (3.9), Combined 64.2 (4.4). CO₂ emissions: 114g/km. Standard EU Test figures for comparative purposes and may not reflect real driving results. For more information please visit audi.co.uk/a6